UNIVERSITY OF BIRMINGHAM
University of Birmingham Research at Birmingham

Optimal primary care management of clinical osteoarthritis and joint pain in older people:

Hay, Elaine; Dziedzic, Krysia; Foster, Nadine; Peat, George; van der Windt, Danielle; Bartlam, Bernadette; Blagojevic-Bucknall, Milisa; Edwards, John; Healey, Emma; Holden, Melanie; Hughes, Rhian; Jinks, Claire; Jordan, Kelvin; Jowett, Sue; Lewis, Martyn; Mallen, Christian; Morden, Andrew; Nicholls, Elaine; Ong, Bie Nio; Porcheret, Mark

DOI:
10.3310/pgfar06040

License:
Other (please specify with Rights Statement)

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (Harvard):
Hay, E, Dziedzic, K, Foster, N, Peat, G, van der Windt, D, Bartlam, B, Blagojevic-Bucknall, M, Edwards, J, Healey, E, Holden, M, Hughes, R, Jinks, C, Jordan, K, Jowett, S, Lewis, M, Mallen, C, Morden, A, Nicholls, E, Ong, BN, Porcheret, M, Kigozi, J, Oppong, R, Paskins, Z & Croft, P 2018, 'Optimal primary care management of clinical osteoarthritis and joint pain in older people: a mixed-methods programme of systematic reviews, observational and qualitative studies, and randomised controlled trials', *Programme Grants for Applied Research*, vol. 6, no. 4. https://doi.org/10.3310/pgfar06040

Link to publication on Research at Birmingham portal

Publisher Rights Statement:
© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay et al. under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

General rights
Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

•Users may freely distribute the URL that is used to identify this publication.
•Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
•User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
•Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy
While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

Download date: 17. Apr. 2024

PROGRAMME GRANTS FOR APPLIED RESEARCH

VOLUME 6 ISSUE 4 JULY 2018
ISSN 2050-4322

Check for updates

Optimal primary care management of clinical osteoarthritis and joint pain in older people: a mixed-methods programme of systematic reviews, observational and qualitative studies, and randomised controlled trials

Elaine Hay, Krysia Dziedzic, Nadine Foster, George Peat, Danielle van der Windt, Bernadette Bartlam, Milisa Blagojevic-Bucknall, John Edwards, Emma Healey, Melanie Holden, Rhian Hughes, Clare Jinks, Kelvin Jordan, Sue Jowett, Martyn Lewis, Christian Mallen, Andrew Morden, Elaine Nicholls, Bie Nio Ong, Mark Porcheret, Jerome Wulff, Jesse Kigozi, Raymond Oppong, Zoe Paskins and Peter Croft

NHS
National Institute for Health Research

DOI 10.3310/pgfar06040

Optimal primary care management of clinical osteoarthritis and joint pain in older people: a mixed-methods programme of systematic reviews, observational and qualitative studies, and randomised controlled trials

Elaine Hay,[1]* Krysia Dziedzic,[1] Nadine Foster,[1] George Peat,[1] Danielle van der Windt,[1] Bernadette Bartlam,[1] Milisa Blagojevic-Bucknall,[1] John Edwards,[1] Emma Healey,[1] Melanie Holden,[1] Rhian Hughes,[1] Clare Jinks,[1] Kelvin Jordan,[1] Sue Jowett,[1,2] Martyn Lewis,[1] Christian Mallen,[1] Andrew Morden,[1] Elaine Nicholls,[1] Bie Nio Ong,[1] Mark Porcheret,[1] Jerome Wulff,[1] Jesse Kigozi,[1,2] Raymond Oppong,[1,2] Zoe Paskins[1] and Peter Croft[1]

[1]Arthritis Research UK Primary Care Centre, Institute of Primary Care and Health Sciences, Keele University, Keele, UK
[2]Health Economics Unit, University of Birmingham, Birmingham, UK

*Corresponding author

Declared competing interests of authors: Krysia Dziedzic was appointed a National Institute for Health and Care Excellence Fellow during the programme period (2013–16), received an NHS England Regional Innovation Fund award to implement aspects of the programme, and was an invited speaker by the British Health Professionals in Rheumatology to present the results. John Edwards was in receipt of a National Institute for Health Research In-Practice Fellowship (2010–12) to carry out parts of this programme and was an invited speaker by the European League Against Rheumatism to present the results. He is also a general practice contractor and benefits from payments under the Quality and Outcomes Framework of the General Medical Services Contract.

Published July 2018
DOI: 10.3310/pgfar06040

This report should be referenced as follows:

Hay E, Dziedzic K, Foster N, Peat G, van der Windt D, Bartlam B, *et al*. Optimal primary care management of clinical osteoarthritis and joint pain in older people: a mixed-methods programme of systematic reviews, observational and qualitative studies, and randomised controlled trials. *Programme Grants Appl Res* 2018;**6**(4).

Programme Grants for Applied Research

ISSN 2050-4322 (Print)

ISSN 2050-4330 (Online)

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (www.publicationethics.org/).

Editorial contact: journals.library@nihr.ac.uk

The full PGfAR archive is freely available to view online at www.journalslibrary.nihr.ac.uk/pgfar. Print-on-demand copies can be purchased from the report pages of the NIHR Journals Library website: www.journalslibrary.nihr.ac.uk

Criteria for inclusion in the *Programme Grants for Applied Research* journal

Reports are published in *Programme Grants for Applied Research* (PGfAR) if (1) they have resulted from work for the PGfAR programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Programme Grants for Applied Research programme

The Programme Grants for Applied Research (PGfAR) programme, part of the National Institute for Health Research (NIHR), was set up in 2006 to produce independent research findings that will have practical application for the benefit of patients and the NHS in the relatively near future. The Programme is managed by the NIHR Central Commissioning Facility (CCF) with strategic input from the Programme Director.

The programme is a national response mode funding scheme that aims to provide evidence to improve health outcomes in England through promotion of health, prevention of ill health, and optimal disease management (including safety and quality), with particular emphasis on conditions causing significant disease burden.

For more information about the PGfAR programme please visit the website: http://www.nihr.ac.uk/funding/programme-grants-for-applied-research.htm

This report

The research reported in this issue of the journal was funded by PGfAR as project number RP-PG-0407-10386. The contractual start date was in July 2008. The final report began editorial review in March 2015 and was accepted for publication in May 2016. As the funder, the PGfAR programme agreed the research questions and study designs in advance with the investigators. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The PGfAR editors and production house have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the final report document. However, they do not accept liability for damages or losses arising from material published in this report.

This report presents independent research funded by the National Institute for Health Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, CCF, NETSCC, PGfAR or the Department of Health and Social Care. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, NETSCC, the PGfAR programme or the Department of Health and Social Care.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk).

NIHR Journals Library Editor-in-Chief

Professor Tom Walley Director, NIHR Evaluation, Trials and Studies and Director of the EME Programme, UK

NIHR Journals Library Editors

Professor Ken Stein Chair of HTA and EME Editorial Board and Professor of Public Health, University of Exeter Medical School, UK

Professor Andrée Le May Chair of NIHR Journals Library Editorial Group (HS&DR, PGfAR, PHR journals)

Dr Martin Ashton-Key Consultant in Public Health Medicine/Consultant Advisor, NETSCC, UK

Professor Matthias Beck Professor of Management, Cork University Business School, Department of Management and Marketing, University College Cork, Ireland

Dr Tessa Crilly Director, Crystal Blue Consulting Ltd, UK

Dr Eugenia Cronin Senior Scientific Advisor, Wessex Institute, UK

Dr Peter Davidson Director of the NIHR Dissemination Centre, University of Southampton, UK

Ms Tara Lamont Scientific Advisor, NETSCC, UK

Dr Catriona McDaid Senior Research Fellow, York Trials Unit, Department of Health Sciences, University of York, UK

Professor William McGuire Professor of Child Health, Hull York Medical School, University of York, UK

Professor Geoffrey Meads Professor of Wellbeing Research, University of Winchester, UK

Professor John Norrie Chair in Medical Statistics, University of Edinburgh, UK

Professor John Powell Consultant Clinical Adviser, National Institute for Health and Care Excellence (NICE), UK

Professor James Raftery Professor of Health Technology Assessment, Wessex Institute, Faculty of Medicine, University of Southampton, UK

Dr Rob Riemsma Reviews Manager, Kleijnen Systematic Reviews Ltd, UK

Professor Helen Roberts Professor of Child Health Research, UCL Great Ormond Street Institute of Child Health, UK

Professor Jonathan Ross Professor of Sexual Health and HIV, University Hospital Birmingham, UK

Professor Helen Snooks Professor of Health Services Research, Institute of Life Science, College of Medicine, Swansea University, UK

Professor Jim Thornton Professor of Obstetrics and Gynaecology, Faculty of Medicine and Health Sciences, University of Nottingham, UK

Professor Martin Underwood Director, Warwick Clinical Trials Unit, Warwick Medical School, University of Warwick, UK

Please visit the website for a list of editors: www.journalslibrary.nihr.ac.uk/about/editors

Editorial contact: journals.library@nihr.ac.uk

Abstract

Optimal primary care management of clinical osteoarthritis and joint pain in older people: a mixed-methods programme of systematic reviews, observational and qualitative studies, and randomised controlled trials

Elaine Hay,[1]* Krysia Dziedzic,[1] Nadine Foster,[1] George Peat,[1] Danielle van der Windt,[1] Bernadette Bartlam,[1] Milisa Blagojevic-Bucknall,[1] John Edwards,[1] Emma Healey,[1] Melanie Holden,[1] Rhian Hughes,[1] Clare Jinks,[1] Kelvin Jordan,[1] Sue Jowett,[1,2] Martyn Lewis,[1] Christian Mallen,[1] Andrew Morden,[1] Elaine Nicholls,[1] Bie Nio Ong,[1] Mark Porcheret,[1] Jerome Wulff,[1] Jesse Kigozi,[1,2] Raymond Oppong,[1,2] Zoe Paskins[1] and Peter Croft[1]

[1]Arthritis Research UK Primary Care Centre, Institute of Primary Care and Health Sciences, Keele University, Keele, UK
[2]Health Economics Unit, University of Birmingham, Birmingham, UK

*Corresponding author e.m.hay@keele.ac.uk

Background: Osteoarthritis (OA) is the most common long-term condition managed in UK general practice. However, care is suboptimal despite evidence that primary care and community-based interventions can reduce OA pain and disability.

Objectives: The overall aim was to improve primary care management of OA and the health of patients with OA. Four parallel linked workstreams aimed to (1) develop a health economic decision model for estimating the potential for cost-effective delivery of primary care OA interventions to improve population health, (2) develop and evaluate new health-care models for delivery of core treatments and support for self-management among primary care consulters with OA, and to investigate prioritisation and implementation of OA care among the public, patients, doctors, health-care professionals and NHS trusts, (3) determine the effectiveness of strategies to optimise specific components of core OA treatment using the example of exercise and (4) investigate the effect of interventions to tackle barriers to core OA treatment, using the example of comorbid anxiety and depression in persons with OA.

Data sources: The North Staffordshire Osteoarthritis Project database, held by Keele University, was the source of data for secondary analyses in workstream 1.

Methods: Workstream 1 used meta-analysis and synthesis of published evidence about effectiveness of primary care treatments, combined with secondary analysis of existing longitudinal population-based cohort data, to identify predictors of poor long-term outcome (prognostic factors) and design a health economic decision model to estimate cost-effectiveness of different hypothetical strategies for implementing optimal primary care for patients with OA. Workstream 2 used mixed methods to (1) develop and test a 'model OA consultation' for primary care health-care professionals (qualitative interviews, consensus, training and evaluation) and (2) evaluate the combined effect of a computerised 'pop-up' guideline for general practitioners (GPs) in the consultation and implementing the model OA consultation on practice and patient

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

outcomes (parallel group intervention study). Workstream 3 developed and investigated in a randomised controlled trial (RCT) how to optimise the effect of exercise in persons with knee OA by tailoring it to the individual and improving adherence. Workstream 4 developed and investigated in a cluster RCT the extent to which screening patients for comorbid anxiety and depression can improve OA outcomes. Public and patient involvement included proposal development, project steering and analysis. An OA forum involved public, patient, health professional, social care and researcher representatives to debate the results and formulate proposals for wider implementation and dissemination.

Results: This programme provides evidence (1) that economic modelling can be used in OA to extrapolate findings of cost-effectiveness beyond the short-term outcomes of clinical trials, (2) about ways of implementing support for self-management and models of optimal primary care informed by National Institute for Health and Care Excellence recommendations, including the beneficial effects of training in a model OA consultation on GP behaviour and of pop-up screens in GP consultations on the quality of prescribing, (3) against adding enhanced interventions to current effective physiotherapy-led exercise for knee OA and (4) against screening for anxiety and depression in patients with musculoskeletal pain as an addition to current best practice for OA.

Conclusions: Implementation of evidence-based care for patients with OA is feasible in general practice and has an immediate impact on improving the quality of care delivered to patients. However, improved levels of quality of care, changes to current best practice physiotherapy and successful introduction of psychological screening, as achieved by this programme, did not substantially reduce patients' pain and disability. This poses important challenges for clinical practice and OA research.

Limitations: The key limitation in this work is the lack of improvement in patient-reported pain and disability despite clear evidence of enhanced delivery of evidence-based care.

Future work recommendations: (1) New thinking and research is needed into the achievable and desirable long-term goals of care for people with OA, (2) continuing investigation into the resources needed to properly implement clinical guidelines for management of OA as a long-term condition, such as regular monitoring to maintain exercise and physical activity and (3) new research to identify subgroups of patients with OA as a basis for stratified primary care including (i) those with good prognosis who can self-manage with minimal investigation or specialist treatment, (ii) those who will respond to, and benefit from, specific interventions in primary care, such as physiotherapy-led exercise, and (iii) develop research into effective identification and treatment of clinically important anxiety and depression in patients with OA and into the effects of pain management on psychological outcomes in patients with OA.

Trial registration: Current Controlled Trials ISRCTN06984617, ISRCTN93634563 and ISRCTN40721988.

Funding: This project was funded by the National Institute for Health Research (NIHR) Programme Grants for Applied Research Programme and will be published in full in *Programme Grants for Applied Research Programme*; Vol. 6, No. 4. See the NIHR Journals Library website for further project information.

Contents

List of tables xiii

List of figures xvii

List of boxes xix

List of abbreviations xxi

Plain English summary xxiii

Scientific summary xxv

Chapter 1 Introduction 1
Methodological note 2

Chapter 2 Workstream 1: modelling optimal primary care for osteoarthritis 3
Abstract 3
Introduction 3
Study 1: predicting long-term outcome in people with joint pain and osteoarthritis 4
Methods 4
Results 6
Discussion 8
Conclusion 10
Study 2: summarising evidence regarding the effectiveness of core primary care interventions 11
Methods 11
Results 12
Discussion 15
Conclusion 15
Study 3: estimating cost-effectiveness of delivering core primary care management for osteoarthritis 16
Methods 16
Analysis 20
Results 21
Discussion 21
Conclusion 22

Chapter 3 Workstream 2: The MOSAICS studies 25
Abstract 25
Introduction and overview of the MOSAICS studies 26
The framework and component studies of MOSAICS 27
The MOSAICS studies 32
Group 1 studies: background evidence about osteoarthritis in primary care 32
Group 2 studies: developing interventions to support self-management of osteoarthritis in primary care and to monitor the quality of care delivered 38
Group 3 studies: development and evaluation of novel methods to measure quality of primary care for osteoarthritis patients 47
Group 4 studies: clinical effectiveness and cost-effectiveness of the MOSAICS programme to optimise primary care of osteoarthritis 52

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Group 5 studies: qualitative evaluation of the MOSAICS studies 70
Group 6 studies: wider dissemination of the MOSAICS interventions 76
Overall conclusions 78
Summary: the MOSAICS research story 81

Chapter 4 Workstream 3: improving the effectiveness of exercise for osteoarthritis – the BEEP trial **85**
Abstract 85
Introduction 86
Development 87
Interventions development 88
BEEP trial physiotherapy training programme 91
The BEEP pilot study 99
Aims and objectives 105
Methods 105
Design 105
Setting and participants 105
Invitation, recruitment, consent and randomisation 106
Interventions 108
Outcome measures 109
Sources of bias 110
Sample size and power 113
Statistical analysis 113
Linked qualitative interviews 117
Results 118
Participant recruitment and flow 118
Interventions delivered and adverse events 118
Clinical outcomes (primary and secondary) 123
Intervention credibility and exercise adherence outcomes 131
Health economic outcomes 131
Results from the linked qualitative study 140
Discussion 147
Summary of trial aim and key results 147
Potential reasons for trial results 148
Comparison with other trials and systematic reviews 150
Strengths and limitations 151
Implications for clinical practice 151
Unanswered questions and future research 152
Overall conclusions 152

Chapter 5 Workstream 4: screening for distress in persons with osteoarthritis – POST **153**
A cluster randomised controlled trial of an electronic prompt for comorbidity screening in clinical osteoarthritis consultations in primary care with nested qualitative studies 153
Abstract 153
Background 153
Aims 155
Study protocol: methods and design 155
Study design: qualitative studies 164
Results 166
Discussion 182

Chapter 6 Conclusions **187**
The results of the programme 187
Summary of the main findings 187
Main implications for practice 187
Recommendations for future work 188
Potential topics for new conceptual thinking and discussion 188
Recommendations for new research 189
Wider implementation 189
Overview of the workstream populations 190
Background 190
Comparing baseline characteristics of the recruited populations 190
Response to care 194
Outputs from delivery and dissemination of the programme 195
Partnerships and collaborative appointments 195
Challenges and innovation in research delivery 195
Dissemination 197

Acknowledgements **199**

References **213**

Appendix 1 Patient and public involvement in the programme **237**

Appendix 2 MOSAICS supporting documents **241**

Appendix 3 BEEP trial supporting documents **251**

Appendix 4 Related publications **255**

Appendix 5 Other grants linked to this programme **259**

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

List of tables

TABLE 1 Baseline characteristics of the 3563 NorStOP cohort 3-year responders with clinical OA in at least one joint at baseline, stratified for outcome categories of moderate to severe pain and functional limitation at 3 years 7

TABLE 2 Summary of the results of meta-analyses for the four main comparisons 13

TABLE 3 Model parameters 18

TABLE 4 Transition probabilities for pain health states 19

TABLE 5 Costs (price year 2010) 20

TABLE 6 Baseline utility scores for health states 20

TABLE 7 Base-case results 21

TABLE 8 Distribution of number of joint pain sites and mean (SD) age and BMI of persons in different categories defined by number of painful sites 34

TABLE 9 Reported use of NICE core treatments in consulters with and without an OA diagnostic Read code in their medical records 36

TABLE 10 Assessor ranking of videos, ordered by gold standard rating rank 41

TABLE 11 Participant ratings of the GP consultation training and its delivery 42

TABLE 12 Summary statistics for GP Competency Score by time point 42

TABLE 13 Characteristics of the eight general practices participating in the cluster trial 55

TABLE 14 Comparison between intervention and control practices of all consulters firing template, by (1) e-template quality indicators and (2) routinely recorded management 56

TABLE 15 Details of e-template use in 1060 patients recorded as consulting with joint pain or OA in the intervention practices, by whether or not practice nurse-led clinic attendance occurred 57

TABLE 16 Characteristics of eligible patients who have consulted about joint pain or OA and responded to the first post-consultation questionnaire 63

TABLE 17 Self-reported receipt or uptake of items indicating quality of care among 525 eligible patients who have consulted about joint pain or OA and returned a post-consultation questionnaire (OA-QIs questionnaire) 63

TABLE 18 Self-reported use of items of care and self-management among 525 eligible patients who have consulted about joint pain or OA and returned a post-consultation (KNEST) questionnaire 64

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 19 Process measures of implementation of MOSAICS: exercise, walking and physical activity, and enablement 66

TABLE 20 PASE components analysis: intervention vs. control 66

TABLE 21 Primary outcome measure (SF-12 PCS) by study groups including sensitivity analyses 67

TABLE 22 Questionnaire response rate and characteristics of participating physiotherapists 93

TABLE 23 Physiotherapists' confidence in diagnosing and managing older adults with knee pain 93

TABLE 24 Types of advice included for the patient case vignette by participating physiotherapists 94

TABLE 25 Types of exercise for the patient case vignette reported by physiotherapists 95

TABLE 26 Exercise delivery for the patient case vignette 96

TABLE 27 Physiotherapists' agreement with attitude statements about exercise for older adults with knee pain [*n* (%) who largely or totally agreed] 97

TABLE 28 Summary of the BEEP trial interventions 101

TABLE 29 Summary of key themes from qualitative interviews in the BEEP pilot study 103

TABLE 30 Summary of patient outcomes in BEEP pilot study (all intervention arms are combined) 104

TABLE 31 Summary of outcome measures 111

TABLE 32 Summary of health-care resource unit costs 116

TABLE 33 Descriptive characteristics of participants recruited by each recruitment method in the BEEP trial 119

TABLE 34 Baseline characteristics of participants ($n = 514$) 120

TABLE 35 Summary of results for primary outcomes at each time point 125

TABLE 36 Analysis of the primary outcome at the primary end point (i.e. the 6-month follow-up) stratified by level of self-reported exercise adherence at 6-month follow-up 126

TABLE 37 Summary of results for all secondary outcome measures 126

TABLE 38 Treatment credibility and exercise adherence at each follow-up time point 132

TABLE 39 Summary of mean resource use by treatment arm over 18 months 135

TABLE 40 Mean (SD) costs by treatment arm over 18 months 136

TABLE 41 Summary of EQ-5D-3L and QALY outcomes by treatment group over 18 months 137

TABLE 42 Cost–utility analysis for 18 months' follow-up 137

TABLE 43 Sensitivity analysis mean (SD) costs by treatment arm over 18 months 139

TABLE 44 Description of work-related outcomes for participants in paid employment, by treatment group 140

TABLE 45 Summary of characteristics of interviewees 141

TABLE 46 Factors facilitating uptake and maintenance of exercise and general physical activity 144

TABLE 47 Features of the therapeutic alliance in BEEP trial interactions 145

TABLE 48 Barriers to the uptake and maintenance of exercise and general physical activity 146

TABLE 49 Summary of full postal questionnaire content 160

TABLE 50 Content of GP's experience of POST questionnaire 164

TABLE 51 Outline of interview schedule for GPs 166

TABLE 52 Cluster-level baseline characteristics according to study group 167

TABLE 53 Participant baseline characteristics by study group 168

TABLE 54 Evaluation of the primary outcome measure (current pain intensity) 169

TABLE 55 Secondary outcome measures: mean (standard deviation), mean difference and *p*-value of difference, unless otherwise indicated 171

TABLE 56 Acceptability and fidelity of screening from participants' post-consultation questionnaires 175

TABLE 57 Participating GPs' perceived impact of screening questions on the consultation 175

TABLE 58 Summary of responses to request for interview (GPs) 177

TABLE 59 Baseline comparability of study populations 192

TABLE 60 Response to care in three trial populations: mean PASE scores and their change over time 194

TABLE 61 Unit costs (£) of resource use items 244

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 62 Resource use mean (SD) over 12 months (complete cases) 246

TABLE 63 Mean costs (SD) per patient over 12 months (complete cases) for each resource category in OA care 247

TABLE 64 Mean total costs (SD) per patient over 12 months (imputed analysis) 247

TABLE 65 Health outcomes, mean (SD) over 12 months (imputed analysis) 248

TABLE 66 Base-case cost–utility analysis (imputed analysis) 248

TABLE 67 Sensitivity analysis 249

List of figures

FIGURE 1 The PAR and NNT for the most important predictors of moderate to severe pain at 3 years 8

FIGURE 2 The PAR and NNT for the most important predictors of limitation in function at 3 years 9

FIGURE 3 Effect estimates (SMD) for pain outcomes of exercise interventions for OA 14

FIGURE 4 Funnel plot for trials investigating exercise interventions for OA (pain outcomes) 15

FIGURE 5 Markov model structure 17

FIGURE 6 The NICE core recommendations for treatment of OA (redrawn using relevant components from NICE) 26

FIGURE 7 Cluster trial framework and population 28

FIGURE 8 BEEP recruitment flow diagram 107

FIGURE 9 Means and 95% CIs for the primary outcome at each time point 124

FIGURE 10 Base-case cost–utility plane comparing ITE with UC 138

FIGURE 11 Base-case cost–utility plane comparing TEA with UC 138

FIGURE 12 Base-case cost–utility plane comparing TEA and ITE with UC 139

FIGURE 13 Evidence of intervention constructs described in qualitative interviews in all three intervention arms of the BEEP trial 141

FIGURE 14 Flow chart of main study design 156

FIGURE 15 Snapshots of electronic screening template for (a) the intervention arm; and (b) the control arm 158

FIGURE 16 Base-case cost–utility analysis (imputed analysis): CEAC 249

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

List of boxes

BOX 1 Interpretation of indicators 9

BOX 2 'Read coding' of OA-related consultations 29

BOX 3 Interview extract (34:82 C) 181

BOX 4 Interview extract (34:85 C) 182

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

List of abbreviations

1RM	1 repetition maximum
ABC knee	Attitudes and Beliefs Concerning knee pain
AHSN	Academic Health Sciences Network
AMED	Allied and Complementary Medicine Database
ANCOVA	analysis of covariance
APEX	Acupuncture, Physiotherapy and Exercise
AUSCAN	Australian/Canadian Osteoarthritis Hand Index
BEEP	Benefits of Effective Exercise for knee Pain
BMI	body mass index
BNF	*British National Formulary*
CACE	complier average causal effect
CCG	Clinical Commissioning Group
CEAC	cost-effectiveness acceptability curve
CI	confidence interval
CLRN	Clinical Research Network
CRF	case report form
CTU	clinical trials unit
DH	Department of Health
DVD	digital versatile disc
EMIS	Egton Medical Information System
EQ-5D	EuroQol-5 Dimensions
EQ-5D-3L	EuroQol-5 Dimensions, three-level version
EQ-5D-5L	EuroQol-5 Dimensions, five-level version
EULAR	European League Against Rheumatism
FPDI	Foot Pain Disability Index
GAD	Generalised Anxiety Disorder
GAD-2	two-item Generalised Anxiety Disorder Assessment
GAD-7	seven-item Generalised Anxiety Disorder Assessment
GP	general practitioner
ICC	intracluster correlation coefficient
ICECAP-A	ICEpop CAPability measure for Adults
ICER	incremental cost-effectiveness ratio
IPAQ	International Physical Activity Questionnaire
IPCHS	Institute of Primary Care and Health Sciences
IQR	interquartile range
IRR	incidence rate ratio
ITE	individually tailored exercise
JIGSAW	Joint Innovation of Guidelines for Osteoarthritis in the West Midlands
KNEPP	Knee Pain Prevention Project
KNEST	Knee Pain Screening Tool
MOAC	model osteoarthritis consultation
MOSAICS	Managing OSteoArthritis In ConsultationS
MRC	Medical Research Council
MUST	Musculoskeletal pain in Ullensaker Study Osteoarthritis
MUST OA-QI	Musculoskeletal pain in Ullensaker Study Osteoarthritis Cohort Quality Indicator
NICE	National Institute for Health and Care Excellence
NIHR	National Institute for Health Research
NNT	number needed to treat
NorStOP	North Staffordshire Osteoarthritis Projects

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

NPT	normalisation process theory
NRS	numerical rating scale
NSAID	non-steroidal anti-inflammatory drug
OA	osteoarthritis
OA-QI	osteoarthritis quality indicator
OARSI	Osteoarthritis Research Society International
OEE	Outcome Expectations for Exercise
OMERACT	Outcome Measures in Rheumatoid Arthritis Clinical Trials
OR	odds ratio
PAR	population attributable risk
PASE	Physical Activity Scale for the Elderly
PCRN	Primary Care Research Network
PCS	physical component score
PCT	primary care trust
PhD	Doctor of Philosophy
PHQ	Patient Health Questionnaire
PHQ-2	Patient Health Questionnaire 2
PHQ-8	Patient Health Questionnaire 8
PHQ-9	Patient Health Questionnaire 9
POST	Primary care Osteoarthritis Screening Trial
PPI	patient and public involvement
PSC	practitioner self-confidence
PSS	Personal Social Services
PSSRU	Personal Social Services Research Unit
QALY	quality-adjusted life-year
QOF	Quality and Outcomes Framework
QoL	quality of life
RCT	randomised controlled trial
REC	Research Ethics Committee
ROC	receiver operating characteristic
RUG	research user group
SAE	serious adverse event
SD	standard deviation
SEE	Self-Efficacy for Exercise
SF-12	Short Form questionnaire-12 items
SF-36	Short Form questionnaire-36 items
SF-6D	Short Form questionnaire-6 Dimensions
SMART	specific, measurable, achievable, realistic, time related
SMD	standardised mean difference
STAR	Short Telephone Activity Rating
TDF	Theoretical Domains Framework
TEA	targeted exercise adherence
TOPIK	Treatment Options for Pain in the Knee
UC	usual physiotherapy care
WISE	Whole systems Implementing Self-management Engagement
WOMAC	Western Ontario and McMaster Universities Osteoarthritis Index

Plain English summary

Osteoarthritis (OA) is a common painful condition in mid to later life that affects daily activities such as walking and climbing stairs. Research shows that simple things such as providing information about OA and exercise can reduce the symptoms, but many sufferers do not receive this information or get advice and encouragement to put it into practice.

Our programme addressed how to make OA interventions more available and more effective by conducting four large research studies:

1. We developed ways to use existing data sets to investigate how to offer the best possible treatments for OA in ways that are effective for patients as well as being affordable for the NHS.
2. In a large trial, we tested how the best care for OA can be delivered in general practitioner (GP) practices with an OA guidebook and structured support by the GP and practice nurse. We showed that it is possible to improve the quality of OA care and we are now using this information to implement best-quality OA care more widely in primary care in the NHS.
3. We conducted a trial comparing different ways to help people to find the right exercise routine and maintain it over time. We found that more intense exercise programmes did not improve symptoms more than exercise already offered by NHS physiotherapists to people with OA.
4. We investigated detecting and managing concurrent depression and anxiety in people with OA. This did not improve OA symptoms more than usual care in general practice.

Overall, this programme has added valuable information about which primary care treatments are effective and which are not effective for people with OA. It has also provided evidence to guide the direction of future research into improving care for people with OA.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Scientific summary

The National Institute for Health Research (NIHR) programme described in this report concerns people with osteoarthritis (OA) and their primary health care.

Osteoarthritis, a common cause of pain and stiffness in the joints in older adults (adults aged ≥ 45 years), is:

- a major global cause of disability
- the most frequent cause of restricted physical and social activity in older people in the UK
- the most common long-term condition managed in UK general practice
- estimated to become the dominant preventable cause of chronic disability in the UK by the year 2030.

Osteoarthritis is a long-term condition. Relief of pain and the maintenance of active participation in daily work, domestic and social life are the goals for treatment of most patients, rather than complete cure. Dramatic improvements can be achieved by joint replacement surgery, which is conducted in the minority of patients who develop advanced and severe OA.

The context for the programme

This programme was formulated in 2008–9 on the basis of questions raised by evidence from two sources about OA and its treatment. First was the published research evidence base about OA which is considered below. Second was input from, and work carried out with, members of the research user group (RUG) in the Institute of Primary Care and Health Sciences.

1. There was evidence that:

    i. simple primary care interventions can provide short-term relief of pain and restricted activity in persons with OA. This evidence underpinned core guidance on care for patients with OA from the UK National Institute for Health and Care Excellence (NICE) published in 2008 (National Institute for Health and Clinical Excellence. *Osteoarthritis: The Care and Management of Osteoarthritis in Adults*. Clinical Guideline CG59. London: NICE; 2008)
    ii. most people with OA have other health conditions (comorbidities) and if effective care is provided for these other conditions, then the pain and disability of OA improves
    iii. self-management programmes can improve outcomes in patients with long-term conditions, such as irritable bowel syndrome, which, like OA, are managed predominantly in primary care and the community.

2. However, there was also evidence that:

    i. the quality of OA primary care ranks as low compared with other long-term conditions
    ii. there was much variability in the content and quality of primary care offered to patients with OA, and treatments often targeted the diseased joint rather than the person with symptoms
    iii. most persons with OA were not receiving the full range of NICE core treatments for OA, including advice and information
    iv. persons with OA and other long-term conditions want help and support to self-manage their condition
    v. short-term changes in behaviour that benefit persons with OA, such as increased physical activity, are often not maintained in the long term
    vi. comorbidities in persons with long-term painful conditions such as OA are under-recognised in primary care.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

3. There was an absence of evidence and information about important aspects of OA care in the community and primary care, including:
   i. no clear estimate of the potential for population prevention and health-care benefits among older people with joint pain and disability in the UK community. Small effects from simple core treatments applied to large numbers of people with OA could result in substantial population health gains at reasonable cost
   ii. the best way for general practitioners (GPs) and the primary care team to consult with persons with OA to manage their condition and support self-management. There had been few attempts to involve OA patients in shaping what information they wanted or needed and what to do to manage their condition
   iii. no formally developed indicators of the quality of primary OA care such as those that existed for other long-term conditions in the UK.

4. There was evidence that most OA patients and health-care professionals were pessimistic (or at best stoical) about the long-term course of OA and the likelihood of getting effective care until the condition was bad enough to require an operation. Patients often perceived doctors' views to be negative: 'not much to be done' and 'what else can be expected at your age?'.

The conclusion was that OA care in the UK was suboptimal, despite evidence that primary care and community-based interventions can reduce pain and disability. The research gaps identified included the need for development and evaluation of better information for public health, patients and clinicians, and for new approaches to delivering OA care in practice.

The overall ambition of the programme was to provide an evidence base for optimising the primary care delivered to OA patients by:

- filling gaps in the evidence and information base about optimal primary care for persons with OA
- providing evidence about how to close the gap between established evidence of effective and efficient OA care and the effective and efficient delivery of that care to patients in practice
- providing practical evidence-based outputs and resources for primary care to support self-management by patients with OA.

The programme consisted of four workstreams. These workstreams were developed in parallel during the period of the programme, linked but independent.

The general aims were as follows.

Workstream 1: to develop a model for estimating the potential for effectively delivered primary care OA interventions to improve population health.

Workstream 2: to develop and evaluate new approaches and services for better delivery and monitoring of OA primary care – the Managing Osteoarthritis In ConsultationS (MOSAICS) studies.

Workstream 3: to investigate the long-term effect of targeting patients' adherence to physiotherapy-led exercise and physical activity for OA knee – the Benefits of Effective Exercise for knee Pain (BEEP) trial.

Workstream 4: to investigate the effect of screening for two major comorbidities (depression and anxiety) in patients with OA in primary care – the Primary care Osteoarthritis Screening Trial (POST).

Workstream 1

Objective
To develop a model to estimate cost-effective population prevention of pain and disability by interventions in patients with OA in primary care.

Methods
Meta-analysis and synthesis of published evidence about effectiveness of primary care treatments, combined with secondary analysis of existing longitudinal population-based cohort data, to identify predictors of poor long-term outcome (prognostic factors); and creation of a health economic decision model to estimate cost-effectiveness of different hypothetical strategies for implementing optimal primary care for patients with OA.

Results
A hypothetical model of effectively delivered 'stepped care' for OA, comprising good advice and pain relief for all patients with OA and offering more extensive treatment (including supervised exercise) to those with poor response to initial treatment was likely to be more cost-effective than current primary care. However, long-term cost-effectiveness is still in need of an evidence base.

Practical outputs
Practical outputs include a population impact model that can be updated and adapted with new information on prognosis and treatment of OA in the general population. The lack of published research and evidence about a constructive and effective approach for long-term analgesic use by OA patients was highlighted.

Workstream 2: the MOSAICS studies

Objectives

- To develop and evaluate a linked 'model' GP and practice nurse consultation and optimal general practice-based service provision for patients with OA (the 'MOSAICS model-of-care') as support for self-management and as a means to implement NICE core OA guidance (NICE, 2008).
- To develop and evaluate methods to measure and monitor the quality of primary care for patients with OA.
- To determine the clinical effectiveness and cost-effectiveness of implementing NICE core OA guidance through the MOSAICS model-of-care in general practice.

Methods
A suite of studies, using a mixture of methods, within the framework of a cluster randomised controlled trial (RCT). Outcomes were (1) process [to measure implementation of evidence-based care informed by NICE guidance (NICE, 2008)] and (2) clinical (to measure the impact on patient symptoms and activities of daily life).

Results
The NICE core OA guidance (NICE, 2008) can be successfully implemented and monitored in primary care practice and is well received by patients. However, the impacts on pain and disability for primary care OA patients have not been demonstrated.

Practical outputs
Practical outputs developed, delivered and evaluated in the programme included an OA guidebook developed with patients and health-care professionals, a training programme for a 'model OA consultation' by GPs, a practice nurse-led primary care service for OA and a training programme for practice nurses, an OA quality-of-care monitoring tool for use in routine primary care computerised records, and a general practice-based OA service model that is being disseminated.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Workstream 3: the BEEP trial

Objective
To determine whether or not early changes in pain and disability in patients with knee OA from physiotherapy-led exercise and physical activity can be better maintained in the long term by enhanced interventions targeting adherence.

Methods
Individual patient three-arm RCT comparing two enhanced exercise interventions led by physiotherapists with usual physiotherapy care (UC) in patients with knee OA (pilot and full trial phases).

Results
Neither of the two experimental enhancements bettered the clinical effectiveness and cost-effectiveness of UC for pain and disability in patients with OA of the knee.

Practical outputs
Renewed support for ensuring that current best practice for physiotherapy-supported physical activity is available to all patients and evidence to support (1) future research identifying the characteristics of patients who respond to exercise interventions and (2) the importance of the therapeutic relationship in improving long-term change in exercise habits.

Workstream 4: the Primary care Osteoarthritis Screening Trial

Objective
To investigate whether or not GP screening for depression and anxiety among older primary care patients presenting with joint pain or OA is acceptable and is clinically effective and cost-effective in improving pain and disability.

Methods
Pragmatic two-arm cluster RCT (pilot and full trial phases).

Results
Systematically screening older patients who present with joint pain and OA in UK general practice is not effective in reducing pain and disability.

Practical outputs
Renewed support for focusing on the OA symptoms (notably joint pain) with which patients present to the GP. One implication of the trial for future research is the need to investigate the usefulness of screening and identifying comorbid joint pain, depression and anxiety in patients being monitored for other long-term conditions.

Conclusions

The programme has provided evidence:

- about the potential for a hypothetical stepped-care model of primary care for patients with joint pain to have a cost-effective impact on population levels of disabling OA
- about ways to implement practical care and support for self-management of OA in UK general practice
- about how to put optimal primary care for OA informed by NICE recommendations into place in the general practice setting
- about ways to monitor and improve the quality of care for patients with OA in UK general practice

- against an additional effect on pain and disability outcomes of adding enhanced interventions to current effective 'usual' physiotherapy-led exercise interventions for patients with knee OA
- against adding screening to current best practice for OA in UK general practice for comorbid depression and anxiety in older patients presenting with musculoskeletal pain.

However, despite workstream 2 providing evidence of successful implementation of optimal evidence-based primary care practice for patients with OA, as supported and endorsed by patients with OA, clinical outcomes (pain and disability) in the study population did not improve as a result. Furthermore, despite successful introduction of a novel screening programme in workstream 4 for two known predictors of poor long-term outcome in older patients with joint pain (depression and anxiety), this also did not lead to subsequent improvement in pain and disability.

Explanations for this absence of impact on clinical outcomes include:

- the possibility that the total pool of patients aged ≥ 45 years with joint pain during the GP consultation in these studies, and considered to have the potential to benefit from these primary care interventions, contains many patients with a good prognosis unaffected by the interventions
- the possibility that the interventions resulted in too little appropriately intensive evidence-based treatment for patients with joint pain who were at moderate or high risk of persistent pain and disability. This includes, in workstream 2, the lack of physiotherapy-led exercise targeted at these groups and, in workstream 4, the lack of active treatments for patients with anxiety and depression targeted at these groups.

But, workstream 1 has provided the basis for modelling of new evidence to investigate alternative strategies, such as stepped care, for resolving these issues. Workstream 2, despite the lack of change in clinical outcomes, has provided evidence of successful implementation of interventions that had been previously evaluated as clinically effective and cost-effective by NICE. This is now the basis for a wider programme of dissemination and evaluation. Workstream 3 patients were selected to have moderate or more severe pain and there was evidence that intensive interventions designed to improve adherence to exercise had been delivered. Usual-care physiotherapy improves pain and disability, but the two novel interventions developed for the BEEP trial did not achieve improved primary clinical outcomes. There was evidence from both the quantitative and qualitative studies that early gains from physiotherapy are not maintained for more than 3 months after the end of physiotherapy treatment. It is likely that community resources and cultural changes, as well as the improvement of therapeutic relationships, are needed to extend the established effectiveness of physiotherapy-led exercise interventions.

Implications for future research and NHS care for patients with osteoarthritis

The conclusions above raise important questions about the need for new conceptual thinking and discussion on OA management in primary care and about the need for further research.

Potential topics for new conceptual thinking and discussion

1. NICE core guidance in practice:

    i. Clarification of the expected benefits of implementing NICE core guidance for patients with OA as a desirable end in itself (e.g. provision of appropriate information and of advice about exercise and pain relief), regardless of whether or not it demonstrably improves clinical outcomes.
    ii. Critical reflection on whether or not current NICE core guidance, even when implemented in the way that MOSAICS has done, results in big enough shifts in crucial components of patient behaviour to achieve change in clinical outcomes in the short term. An example of this is physical activity when there is evidence from other research that implementation of physiotherapy-guided interventions as

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

core treatment could achieve bigger change, so that adding such resource to the MOSAICS package might deliver effects on pain and disability.
   iii. Following on from (ii) above, reviewing the delivery of NICE core guidance for primary care and the capacity of a GP- and nurse-led service alone to improve clinical outcomes without additional resources such as physiotherapy services to provide individualised and supervised exercise interventions.

2. Continuing debate and critical enquiry about the role, benefits and costs of systematic screening for anxiety and depression in all people with long-term conditions in primary care.
3. The need for new concepts about OA within the research and clinical community:

   i. regarding achievable goals of long-term care for people with this condition, including whether or not it is appropriate to seek more than small short- to medium-term clinical effect sizes in patients' pain and disability in practice. Maintenance of activity and participation, despite continuing pain, may, for example, be a better long-term measure of effect.
   ii. regarding a combined approach to the management of OA as a long-term condition (importance of good information, adequate advice and resources for exercise and physical activity and other core interventions, and identification of individuals at high risk of unfavourable future course and targeted interventions for those most likely to benefit).

Recommendations for new research

1. Outcomes in long-term conditions.
2. Research into new models of long-term care for OA, in the context of other long-term disabling conditions, that focus on the necessary and desirable process and clinical outcomes from patient, clinician and societal perspectives.
3. Stratified care for OA patients.
4. Research to identify subgroups of patients with OA who may, on the basis of combined evidence from previous cohorts, clinical effectiveness and cost-effectiveness studies, and the successful implementation strategies described in this programme, benefit from specific interventions in primary care. This includes identification of patient subgroups such as (i) those with good prognosis who can be supported to self-manage without additional investigation or treatment and (ii) those who will benefit from specific treatments, such as physiotherapy-led exercise.
5. Exercise and physical activity levels in people with long-term conditions:

   i. research to identify new approaches to improving long-term adherence to exercise and physical activity among patients with OA such as regular monitoring.

6. Depression and anxiety in people with OA:

   i. research into more efficient and effective ways of identifying and treating clinically important levels of anxiety and depression in patients with OA.
   ii. research into the effects of pain management on psychological outcomes in patients with OA.

Trial registration

The trials in this programme are registered as ISRCTN06984617, ISRCTN93634563 and ISRCTN40721988.

Funding

Funding for this study was provided by the Programme Grants for Applied Research programme of the National Institute for Health Research.

Chapter 1 Introduction

The National Institute for Health Research (NIHR) programme described in this report concerns people with osteoarthritis (OA) and their primary health care.

Osteoarthritis is a condition of pain and stiffness in the joints that is common among older adults (adults aged ≥ 45 years) and constitutes a major portion of the global causes of disability.[1] It is the most frequent cause of restricted physical and social activity in older people in the UK,[2] the most common long-term condition managed in UK general practice[3,4] and estimated to become the dominant preventable cause of chronic disability in the UK by the year 2030.[5]

Osteoarthritis is a long-term condition. Relief of pain and the maintenance of active participation in daily work, domestic and social life are the goals for treatment and management of most patients, rather than complete cure.[6] However, dramatic improvements can be achieved by joint replacement surgery, which is conducted in the minority of patients who develop advanced and severe OA.[7]

At the time that this programme was formulated (2008–9), there was evidence that simple interventions that were available in primary care could provide short-term relief of pain and restricted activity in persons with OA.[8–14] This evidence provided the basis and support for core guidance on care for patients with OA by the UK National Institute for Health and Care Excellence (NICE), which was published in 2008.[15] It was also apparent that most people with OA have other health conditions (comorbidities)[16,17] and, if effective care is provided for these other conditions, the pain and disability of OA improves.[18] We also knew that self-management programmes could improve outcomes in patients with long-term conditions (such as irritable bowel syndrome) which, like OA, are managed predominantly in primary care and the community.[19–21] However, there was also evidence that the quality of OA care ranked lower than care for other long-term conditions managed in primary care[22,23] and there was much variability in the content and quality of primary care offered to patients with OA.[24] For example, most persons with OA were not receiving the full range of NICE core treatments for OA,[24] including advice and information regarded by NICE as a core requirement.[15] Patients with OA told us that they wanted help and support to self-manage their condition,[6,22,25–28] because short-term changes in behaviour that benefit persons with OA, such as increased physical activity, are often not maintained in the long term,[29–34] and their comorbidities were often under-recognised in primary care and the community.[35–39]

Our interpretation of the literature at that time highlighted an absence of evidence and information about key aspects of OA care in the community and primary care, including no clear estimate of the potential for prevention of OA in the population and health-care benefits among older people with joint pain and disability in the UK community. It is likely that small effects from simple core treatments applied to large numbers of people with OA could result in substantial population health gains at reasonable cost.[40] Although evidence about the effectiveness of certain non-pharmacological interventions was encouraging, we did not know the best way for general practitioners (GPs) and the primary care team to consult with persons with OA to deliver these interventions and support self-management, or how to measure quality of OA primary care.[23] Prior to our programme, there had been few attempts to involve OA patients in shaping what information they wanted, or needed, and what to do to manage their condition[24] (see *Appendix 1*). Lastly, there was evidence that most OA patients and health-care professionals were pessimistic (or, at best, stoical) about the long-term course of OA and the likelihood of getting effective care until the condition was bad enough to require an operation.[22,41,42] Patients perceived doctors' views to often be negative: 'not much to be done' and 'what else can be expected at your age?'.[41–44]

The conclusion was that OA care in the UK was not as good as it could be, despite evidence that primary care and community-based interventions can reduce pain and disability. The research gaps identified included the need for development and evaluation of better information for public health, patients and

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

clinicians, and for new approaches to delivering OA care in practice. The ambition of our programme was to address these gaps.

Methodological note

The studies in this programme were developed in parallel and not in series. This had two methodological implications. First, the development of a theoretical model for estimating the cost-effectiveness of different primary care approaches to OA management (study 3, workstream 1) could not draw on the results of the other workstreams. However, it has been designed so that it can be populated with new data in the future. Second, EuroQol-5 Dimensions, five-level version (EQ-5D-5L) was not available at the time of the development of the protocols for the studies and this explains the use of the earlier EuroQol-5 Dimensions, three-level version (EQ-5D-3L) in this programme.

Chapter 2 Workstream 1: modelling optimal primary care for osteoarthritis

Parts of this chapter have been reproduced from Wulff.[59]

Abstract

Background: The cost-effectiveness of primary care approaches for patients with OA in primary care is not known.

Objectives: (1) Identify key predictors of 3-year outcomes of pain and function in a population-based sample of people aged ≥ 50 years with joint pain and OA, (2) summarise evidence regarding the effectiveness of core primary care interventions for OA and (3) design a decision model to estimate the cost-effectiveness of implementing evidence-based, core primary care management for OA.

Methods: Secondary analysis of population cohort data to identify predictors. An evidence synthesis and meta-analysis to summarise evidence of the effectiveness of four core primary care interventions [advice and information, simple analgesics, topical non-steroidal anti-inflammatory drugs (NSAIDs), and exercise] for OA. The effect estimates derived from the evidence synthesis were used to populate an economic decision model.

Results: The prediction models showed that, in this population, the strongest predictors of future pain and functional limitation include baseline pain, function, physical activity, general health, obesity and socioeconomic indicators. The models showed good internal validity, but needed to be further developed to account for generalised pain in people with OA and to investigate the predictive value of modifiable prognostic factors (including physical activity and obesity) to identify groups that benefit from specific treatments. The meta-analysis showed statistically significant small to moderate improvements in pain and function from advice and information, topical NSAIDs and exercise interventions compared with controls, while simple analgesia failed to demonstrate statistically significant improvements in either pain or function. The decision model examined the cost-effectiveness of two hypothetical approaches to delivering primary care interventions for hip and knee OA (stepped care and 'one-stop-shop' care) compared with current primary care. The hypothetical stepped-care intervention was the most cost-effective, dominating 'one-stop-shop' and current care, and this result was robust to the sensitivity analyses conducted.

Conclusion: This work provides a platform for modelling the long-term cost-effectiveness of interventions for OA. The model will require further development, including use of more accurate data on the short-term course of OA pain to provide better estimates and the incorporation of evidence on hand and foot OA.

Introduction

Reliable information regarding the likely course of symptoms and factors predicting poor long-term outcomes in people with joint pain and OA is important to inform public health about the burden of these conditions in local communities and, linked with evidence about treatment efficacy and effectiveness, to estimate for policy-makers the probable costs and benefits of health interventions. Such information would provide evidence to inform the potential usefulness of identifying subgroups of persons with OA at higher risk of poor outcomes who may preferentially benefit from targeted timely interventions.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Previous studies have developed prediction models for OA, but most were developed to predict the onset of OA[45,46] or the outcome of a specific treatment,[47,48] or were focused on a specific joint – usually the knee.[49–52] Most people with OA have multiple joint pains[53–55] and our aim was to develop prognostic models and estimate the cost-effectiveness of offering core treatments for OA regardless of the site of pain.

Modelling studies help to identify desirable resource shifts and guide public health and health-care policy.[56] In estimating potential health gains and reductions in disability in relevant subgroups of patients and estimating the costs associated with these gains, policy-makers can determine the gaps between current treatment and optimal management for OA and quantify the need to invest in services and research.[57] Few modelling studies in OA have been carried out in population-based or primary care cohorts[58] and there are not many studies that have used longitudinal data on the long-term outcomes of pain and function.[57] We aimed to address and avoid these limitations by using data from a large, prospective population-based cohort with linked morbidity records from health care in order to estimate the probability and key predictors of unfavourable long-term outcome in people with OA, and the cost-effectiveness of implementing core primary care management for OA.

The objectives of this workstream were to (1) identify key predictors of long-term (3-year) pain and limitation in function in a population-based sample of people with joint pain and OA, (2) synthesise evidence regarding the effectiveness of core primary care interventions and (3) estimate the cost-effectiveness of implementing evidence-based, core primary care management for OA. The study also aimed to provide a basic model as a resource for future analyses of the cost-effectiveness of primary care interventions for joint pain and OA, including novel approaches such as those under investigation in the other workstreams of this programme. The population-based cohort (used in study 1) and the evidence synthesis (study 2) provided data to support the design of this basic health economic model (study 3).

The work described in this chapter was undertaken in the framework of a Doctor of Philosophy (PhD).[59]

The cohort and modelling studies in this workstream were based on an existing population-based cohort [the North Staffordshire Osteoarthritis Projects (NorStOP) study], which included people aged ≥ 50 years. This age criterion contrasts with the criterion of people aged ≥ 45 years who took part in workstreams 2–4 of the programme. The high prevalence of joint pain found in NorStOP was one reason for lowering the minimum age for these other workstreams in order to include interventions early in the course of OA.

Study 1: predicting long-term outcome in people with joint pain and osteoarthritis

This section describes the methods and results of a prognostic study aiming to identify predictors of future long-term (3-year) risk of pain and functional limitation in a population-based sample of older adults with joint pain.

Methods

Design and study population

Data were used from the NorStOP study, a series of population-based prospective cohorts of adults aged ≥ 50 years who were registered with one of eight general practices in North Staffordshire, UK.[60] These cohorts were originally set up through Medical Research Council (MRC) programme funding and NIHR primary care centre funding, and continue to be a resource for researchers from inside and outside Keele University. A two-stage mailing strategy was used, comprising a health survey sent to all persons in the sampling frame and a subsequent regional pain questionnaire sent to those who had responded to the health survey, given consent to be contacted again and indicated they had experienced pain in the hip, knee, hand or foot in the previous year. Participants received similar follow-up questionnaires at 3 years. Ethics approval was obtained from the North Staffordshire Local Research Ethics Committee (NS-LREC

1351 and 1430). This analysis includes all responders who reported hand, hip, knee or foot joint pain at baseline that had lasted ≥ 3 months over the previous 12 months and who returned the 3-year follow-up survey.

Outcome measures

Outcome measures for the prognostic models were (1) moderate or more severe pain in at least one joint site and (2) limitations in physical function in accordance with the Outcome Measures in Rheumatoid Arthritis Clinical Trials (OMERACT) consensus statement on key outcome measures for OA.[61]

The pain subscales of the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) (range 0–20),[62] Australian/Canadian Osteoarthritis Hand Index (AUSCAN) (range 0–20)[63] and Foot Pain Disability Index (FPDI) scores (range –3.32 to 3.33)[64,65] were used to measure pain related to the hip/knee, hand and foot, respectively. All three questionnaires have been validated in relevant populations. To define a dichotomous measure of at least moderate pain, cut-off points for increased scores on the WOMAC, AUSCAN or FPDI scores were selected using receiver operating characteristic (ROC) curve analysis with a 0- to 10-point numerical rating scale (NRS) for pain with a validated cut-off point of 5 as the anchor (0–4 mild pain, 5–10 moderate to severe pain).[66] Based on maximal area under the ROC curve, the cut-off points for at least moderate pain obtained for WOMAC pain were 6 and 5 for the hip and knee, respectively, 9 for AUSCAN hand pain, and –0.479 for FPDI foot pain. Participants with scores below any of the relevant cut-off points at the 3-year follow-up were classified as having 'no or mild pain', while those with scores equal to or above the cut-off point were classified as having 'moderate to severe pain'.

Limitations in physical function were assessed with the 10-item physical functioning subscale of the Short Form questionnaire-36 items (SF-36) (range 0–100, higher scores indicating better physical functioning),[67] which was completed by all participants regardless of the location of pain. The scale was converted into a binary variable based on its median score of 55, as its distribution was heavily skewed. Respondents with a score of less than the median value were classified as having important functional limitation.

Potential predictors

The baseline health survey included questions on physical functioning, joint pain, participation restriction, general health, sociodemographic factors, lifestyle characteristics, comorbidity, medication use and psychosocial factors (including illness perceptions, anxiety and depression). Full details of the variables and their descriptions have been reported elsewhere.[59] These variables were all considered as potential predictors in the prognostic analysis as they have been shown to be associated with symptoms of OA or are considered to be potential predictors of outcome in pain conditions more generally. All variables with prevalence of < 10% or > 90% or with fewer than 30 participants giving a positive response were excluded from the analysis to facilitate successful convergence of the models[68] and optimal discrimination between people with favourable or unfavourable outcome.

Statistical analysis

Model development and internal validation

Poisson regression was used to develop prognostic models, as it estimates the true risk as incidence rate ratio (IRR) for each predictor. The number of events (i.e. those with at least moderate pain or functional limitation at follow-up) was high and consequently the odds ratios (ORs) from logistic regression would not approximate true underlying risk, which was required to estimate population attributable risk (PAR) and number needed to treat (NNT) (see *Selection of the most important predictors based on population attributable risk and number needed to treat*). Robust variance estimation was used to provide accurate effect estimates and their uncertainty.[69] The models were constructed using a backward stepwise variable selection procedure, retaining only those variables with a statistically significant association with outcome ($p < 0.05$) based on the likelihood ratio test. The *c*-statistic was used as a measure of the predictive performance of the models. Bootstrapping (500 samples) was subsequently used to adjust performance estimates for optimism.[70]

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Selection of the most important predictors based on population attributable risk and number needed to treat

Epidemiological parameters (PAR and NNT) were calculated for each predictor to allow for better understanding of the potential impact of individual prognostic factors and to support the selection of the most relevant predictors for identification of vulnerable subgroups of people with joint pain.[71]

The PAR and NNT were estimated for each predictor selected in the models.

The PAR represents the proportion of the risk of poor outcome in the whole population that is explained by a particular predictor. It provides an indication of maximum achievable population health gain if exposure to that predictor was completely eliminated by successful intervention. PAR, adjusted for other predictors in the model, was calculated using the Stata® version 12 (StataCorp LP, College Station, TX, USA) command aflogit, based on a formula proposed by Greenland and Drescher.[72]

The NNT represents the number of people needed to be treated to prevent one additional person from suffering from an unfavourable outcome. It was calculated as the inverse of the absolute risk difference between the exposed and unexposed groups. NNT is usually presented as an estimate of treatment effect.

In this observational study, following the procedure proposed by Smit *et al.*,[71] PAR and NNT were used as impact measures in order to select a limited set of predictors based on high effect size (IRR), high PAR and low NNT. The predictive performance of this key set of predictors was compared with the performance of the original Poisson models.

Non-response and missing values

A non-response analysis has previously been carried out for the NorStOP cohorts by comparing age and sex distribution of responders and non-responders to the baseline questionnaire.[73] We also compared demographic characteristics, pain, physical function, mental health and obesity between participants who completed the 3-year follow-up questionnaire and those dropping out of the study. Furthermore, multiple imputations by chained equation were used to generate data for predictor variables with > 3% missing values. Content and performance of prediction models based on the imputed data sets were similar to that of the complete case analysis; hence, only results of complete cases have been reported here. All analyses were performed using Stata version 11 (StataCorp LP, College Station, TX, USA).

Results

Response

In the original NorStOP cohorts, a total of 26,705 people aged ≥ 50 years were identified from eight general practices and posted the health survey. A total of 18,474 participants responded (estimated response 71.8% after adjustment for age and sex). Among this group of responders were 10,057 persons who reported joint pain in the previous year and gave their consent for further contact. This group was sent a second questionnaire at baseline (the regional pain questionnaire) and a follow-up questionnaire 3 years later.

The cohort for the current analysis, which is described here and was supported by the NIHR programme grant, was the subgroup of 3563 responders at the 3-year follow-up (57.6% of all 3-year regional pain questionnaire responders) who responded at 3 years and had reported pain in at least one joint site at baseline for a duration of ≥ 3 months over the 12 months prior to the baseline survey.

Characteristics of participants

Table 1 presents demographic characteristics of the sample, stratified according to the presence of severe pain and functional limitation at the 3-year follow-up. The mean age of participants was 64 years (range 50–93 years) and 60% were female. Non-responders at baseline and at the 3-year follow-up were slightly younger and more often male. At the 3-year follow-up, more than two-thirds (71%) of the participants

TABLE 1 Baseline characteristics of the 3563 NorStOP cohort 3-year responders with clinical OA in at least one joint at baseline, stratified for outcome categories of moderate to severe pain and functional limitation at 3 years

	Pain at 3 years, *n* (%)		Functional limitation at 3 years, *n* (%)	
	Mild or no pain	Moderate to severe pain	Good function	Poor function
Variables	*N* = 1019 (29%)	*N* = 2544 (71%)	*N* = 2028 (57%)	*N* = 1535 (43%)
Age group (years)				
50–64	556 (55)	1306 (51)	1221 (60)	641 (42)
65–74	322 (32)	852 (34)	617 (30)	557 (36)
≥ 75	141 (13)	386 (15)	190 (10)	337 (22)
Median (IQR)	63 (56–70)	64 (57–71)	62 (56–68)	66 (59–74)
Female sex	567 (56)	1556 (61)	1175 (58)	948 (62)
Married	788 (78)	1810 (72)	1575 (78)	1023 (67)
Employment status				
Employed	314 (32)	584 (24)	729 (37)	169 (11)
Retired	538 (54)	1352 (55)	965 (48)	925 (63)
Unemployed	145 (14)	525 (21)	292 (15)	378 (26)
BMI				
Normal weight (20–24.9 kg/m^2)	383 (39)	728 (29)	736 (38)	375 (26)
Overweight (25–29.9 kg/m^2)	452 (47)	1052 (43)	885 (45)	619 (42)
Obese (≥ 30 kg/m^2)	137 (14)	678 (28)	341 (17)	474 (32)

BMI, body mass index; IQR, interquartile range.
The denominator number used for calculating the column percentages for some categories are different to the column total shown because of missing data for that item.

reported severe pain, while 43% had poor function. The distribution of pain, physical function, obesity, anxiety and depression scores were similar for individuals responding at follow-up and those who dropped out from the study at 3 years (data not shown).

Predictors of poor outcome at three years

Multivariable Poisson regression identified a large number of baseline predictors that were statistically significantly associated with increased risk of moderate to severe pain at 3 years, including variables related to pain severity, poor physical function, socioeconomic factors, obesity, anxiety, hypertension and little use of pain medication or other remedies. Baseline predictors associated with increased risk of functional limitation at 3 years included several measures of poor function, being unemployed or retired from work, lower levels of activity, fewer years in education, joint pain severity, obesity, hypertension and poor health perceptions. Predictive performance was good for both models with *c*-statistics adjusted for optimism of 0.83 [95% confidence interval (CI) 0.81 to 0.85] for predicting moderate to severe pain and 0.91 (95% CI 0.90 to 0.92) for predicting limitation in function at 3 years.

Selection of important predictors

Both models included a large number of predictors, which limited their clinical utility. In this study, the most relevant predictors for severe pain and poor functional limitation at 3 years were selected using the rule employed by Smit *et al.*[71] to select subgroups of participants at increased risk of developing anxiety in later life. The rule selects predictors based on large effect size (IRR), high PAR and low NNT, which have the

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

ability to select factors for which the highest possible health benefit (IRR and PAR) and the lowest possible effort and cost (NNT) can be achieved if interventions are completely successful. In addition, the predictive performance of the set of predictors had to be comparable to that of the original Poisson regression models.

Based on high effect size (IRR), high PAR and low NNT, the following prognostic factors were selected as the most relevant predictors of moderate to severe pain at 3 years: presence of knee pain in the last year (IRR 1.31, 95% CI 1.19 to 1.43), high WOMAC knee pain score (IRR 1.15, 95% CI 1.06 to 1.24), poor physical function (IRR 1.14, 95% CI 1.07 to 1.22), hand pain in the last year (IRR 1.12, 95% CI 1.05 to 1.20), not attended further full-time education after secondary school education has been completed (IRR 1.12, 95% CI 1.01 to 1.23) and obesity (95% CI 1.11, 1.05 to 1.17). The predictive performance (apparent *c*-statistic) of this reduced set of predictors was 0.75 (95% CI 0.73 to 0.77), indicating some reduction in performance for the reduced pain model. The adjusted PAR estimates for these factors were 17%, 9%, 11%, 7%, 9% and 2%, respectively, while their (unadjusted) NNT estimates were $n = 5$, $n = 6$, $n = 4$, $n = 7$, $n = 9$ and $n = 6$, respectively (*Figure 1*).

Using the same procedure for selecting predictors, poor physical function (subscale SF-36, IRR 2.48, 95% CI 2.02 to 3.05), poor physical health [SF-12 physical component score (PCS), IRR 1.44, 95% CI 1.24 to 1.67], being retired from work (IRR 1.39, 95% CI 1.18 to 1.64), reporting reduction in time of or change in activities in the past year (IRR 1.31, 95% CI 1.01 to 1.70), with IRR estimates shown separately for 'much' and 'little' reduction in *Figure 2* and not going out for long walks (> 2 miles on any day) (IRR 1.28, 95% CI 1.10 to 1.48) were identified as the most important predictors for functional limitation (see *Figure 2*). Their adjusted PAR estimates were 48%, 21%, 19%, 8%, 13% and 15%, respectively, with unadjusted NNT estimates of $n = 2$, $n = 2$, $n = 3$, $n = 3$, $n = 2$ and $n = 3$, respectively (see *Figure 2*). The predictive performance (apparent statistic) of this reduced model was 0.71 (95% CI 0.69 to 0.73), indicating a considerable reduction in predictive performance when compared with the full Poisson model.

Discussion

This study identified key baseline predictors of moderate to severe pain and functional limitation at 3 years in a community sample of older people with joint pain or OA. The predictive performance was good for the full prediction models but lower (as expected) for the reduced models, especially for the model predicting functional limitation. Key predictors of poor outcome included high baseline levels of pain and poor function, poor general physical health, low levels of physical activity, indicators of lower socioeconomic status and obesity.

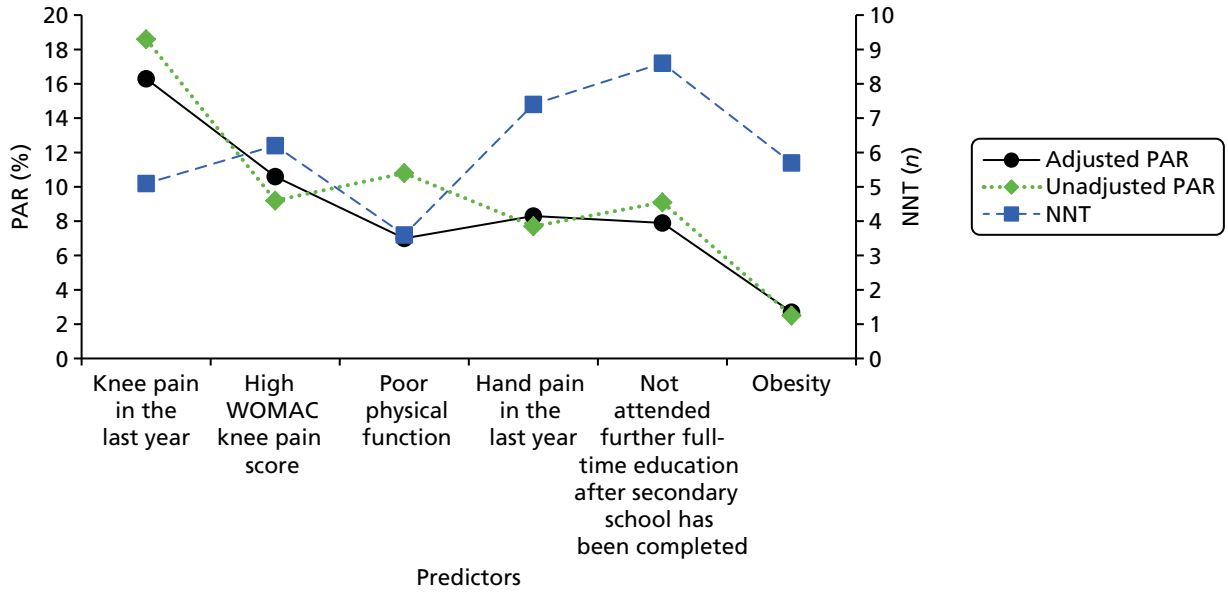

FIGURE 1 The PAR and NNT for the most important predictors of moderate to severe pain at 3 years.

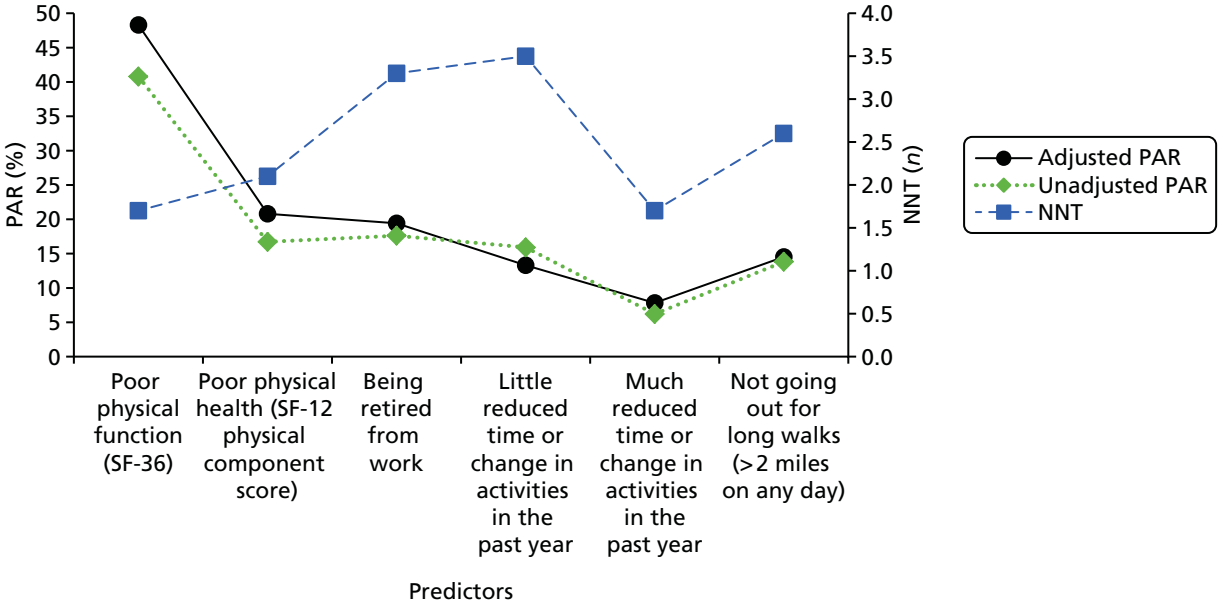

FIGURE 2 The PAR and NNT for the most important predictors of limitation in function at 3 years.

Our data set provided the opportunity to derive prognostic models for long-term outcomes of joint pain in a community sample. The study identified a mix of potentially modifiable and non-modifiable predictors, taking into account a wide range of factors. Selection of predictors for the reduced, more feasible models was not only based on the strength of association of predictors with outcome (IRR), but also on the PAR and NNT associated with these predictors. These indicators take the prevalence of the predictor in the sample into account, providing a measure of impact and facilitating interpretation of findings (*Box 1*).

Although these indicators are useful for presenting findings, and can serve as useful indicators for planning health-care resources and identifying vulnerable subgroups, they have to be interpreted with caution. The PAR assumes there are no other competing risks and, when the PAR is interpreted as an indicator of maximum achievable health gain, the assumption is that predictors are causally associated with outcome and can be optimally addressed by an effective intervention.[74]

This study uniquely derived outcome prediction models for OA regardless of the joints involved. We felt this to be important as most people with OA have multiple pains and may consult with pain and problems at different joints. The onset of multiple joint pain has been shown to be associated with more frequent or severe pain at follow-up and larger increases in locomotor disability.[75,76] The presence of generalised pain is likely to influence the outcome of treatment and should be taken into account when managing musculoskeletal conditions such as OA.[77]

BOX 1 Interpretation of indicators

For example: the adjusted PAR of 17% for presence of knee pain at baseline (IRR = 1.33) indicates that 17% of the risk of severe pain 3 years later in the population of all people who start off with pain in at least one joint (adjusted for known confounders) is explained by having pain located specifically in the knee.

The NNT of $n = 5$ indicates that were baseline knee pain to be successfully treated or prevented in five individuals in this population, this would prevent one additional person having long-term severe pain 3 years later.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Our analysis started with a large number of potential predictors. The shortcomings of using many variables in a stepwise model include a lack of stability of the selected set of predictors and potential bias in the estimation of regression coefficients due to multiple testing.[78] However, testing of our performance estimates using bootstrapping showed limited optimism, indicating adequate internal validity of our models.

Approximately 29% of the total sample did not respond to the baseline health survey. A comparison of non-responders with responders at baseline showed non-responders to be slightly younger and more often male. Therefore, they may also differ with respect to levels of pain, functional limitation or other characteristics. However, this is not likely to greatly influence associations between baseline predictors and 3-year outcomes of pain and function in our subsample of participants with joint pain at baseline. More importantly, approximately one-third of participants did not respond to the follow-up questionnaires at 3 years, mainly because they did not provide consent for further contact or declined to continue with the study. The non-response analysis showed a slightly different age and sex distribution in responders, but very similar baseline scores for other sociodemographic variables and for pain, function, and physical and mental health, limiting the risk of bias when identifying key predictors of long-term pain and functional limitation in this study. A recent study of the potential effects of attrition in the NorStOP cohorts confirmed that there was little evidence that responders at follow-up points represented any further selection bias to that present at baseline.[79]

Conclusion

In a population sample of older people with symptoms of joint pain that is probably attributable to OA, the strongest predictors of moderate to severe pain and functional limitation 3 years later are baseline measures of:

- location and severity of pain
- physical function
- physical activity
- general health
- obesity
- socioeconomic indicators.

Outputs

1. These findings were used as the basis for the modelling of primary care strategies for population health gain in OA (see *Study 3: estimating cost-effectiveness of delivering core primary care management for osteoarthritis*).
2. One potential implication of the findings for the primary care of OA is that improving the most severely affected patients will be needed in order to shift population levels of pain and disability in the long term. However, the analysis presented here was based on a general population sample, which may over-represent prevalent cases when compared with a sample of primary care consulters, and included some characteristics that may not be modifiable, at least in the short term. Future work will need to take better account of the presence of generalised pain in people with joint pain or OA and investigate the specific predictive value of modifiable prognostic factors (including physical activity and obesity) to identify risk groups and individuals for targeting treatment, such as the interventions investigated in the Managing OSteoArthritis In ConsultationS (MOSAICS) studies and Benefits of Effective Exercise for knee Pain (BEEP) trials (see *Chapters 3* and *4*). The second output of this programme is, therefore, a development of the analysis based on a more selective category ('joint pain that interferes with daily life') as the primary outcome, designed to reflect a more severely affected population. Predictors will focus on modifiable prognostic factors for poor outcome.

Study 2: summarising evidence regarding the effectiveness of core primary care interventions

An evidence synthesis and meta-analysis was conducted to summarise available evidence regarding the effectiveness of primary care interventions recommended by NICE for patients with OA (i.e. advice and information, paracetamol, topical NSAIDs and exercise). The aim of the evidence synthesis was to obtain effect estimates that can be used to populate the health economic decision model presented in study 3 (see *Study 3: estimating cost-effectiveness of delivering core primary care management for osteoarthritis*). Interventions to lose weight were not included in this analysis as the decision model did not solely concern overweight or obese patients.

Methods

There are numerous published systematic reviews of the effectiveness of conservative treatments for OA, which made it possible to design an efficient search strategy, identifying relevant randomised trials first from existing systematic reviews and then by an updated search of trials not yet included in the reviews. As this work was conducted at the start of the funding period for this *Programme Grants for Applied Research* report, this evidence synthesis included trials published until August 2010.

Search strategy

Systematic reviews and meta-analyses were searched in MEDLINE and The Cochrane Library from January 1990 to August 2010. Subsequently, an additional search in MEDLINE, covering the period from the year 2000 to August 2010, was conducted to identify individual trials that were not yet included in reviews. The search terms were developed in consultation with an information specialist and were based on the following (exploded) medical subject heading terms: OA, family practice, general practice, primary health care, community health services, ambulatory care, non-steroidal anti-inflammatory agents, exercise therapy. Reference lists from all retrieved systematic reviews and randomised controlled trials (RCTs) were checked to identify additional potentially eligible RCTs.

Selection of trials

Included in the evidence synthesis were full reports of RCTs, published in English, that evaluated the effectiveness of advice or information regarding self-management approaches and of paracetamol, topical NSAIDs and exercise in patients diagnosed (radiographically or symptomatically) with OA at one or more joint sites (hand, hip, knee or foot). Given the focus of this study on primary care, RCTs were only selected if they had been conducted in a primary care or direct access setting. One reviewer (JW) scored eligibility of all identified publications, with an additional reviewer (DvdW, MB or SJ) judging eligibility of all potentially relevant studies.

Data extraction

Data were extracted from RCTs included in this review to enable description of setting, study design, trial population, interventions and outcome measures. Estimates of absolute mean or changes in mean scores after treatment and their respective standard deviations (SD) were extracted and used to calculate standardised mean differences (SMDs) for outcomes of pain and functional limitation.[80] When relevant reports were available, effect estimates were calculated based on intention to treat analysis. Effect estimates of 0.2, 0.5 and 0.8 were described as small, moderate and large, respectively.[81]

Assessment of risk of bias

Risk of bias of RCTs was assessed using The Cochrane Risk of Bias tool.[82]

The risk of bias domains considered most important in this analysis were adequacy of randomisation, concealment allocation, blinding of outcome assessors and loss to follow-up. These were scored by JW as high risk, low risk or unclear risk of bias. Sensitivity analyses were performed, omitting trials with high risk of bias in at least one of the domains considered.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Meta-analysis

Meta-analysis focused on comparisons that were of interest to the design of the decision model (see *Study 3: estimating cost-effectiveness of delivering core primary care management for osteoarthritis*): (1) advice and information versus no treatment, (2) simple analgesics versus advice/placebo/no treatment, (3) topical NSAIDs versus advice/placebo/no treatment and (4) exercise versus advice/simple analgesics/no treatment. Pooled effect estimates for each comparison were calculated using a fixed- or random-effects model depending on the extent of heterogeneity present. Fixed-effects models were used when the I^2 estimate was $\leq 50\%$, otherwise random-effects models were used. Depending on the data presented in the original trial reports, SMDs for pain and functional limitation were derived using either mean change or the final outcome score at the end of treatment. This does not pose a problem in meta-analysis as both scores are considered to be addressing the same underlying intervention effect in RCTs.[83] Differences in mean final scores will be the same as differences in mean change scores if randomisation has been successful and baseline values of outcome measures are similar.

Funnel plots and Egger's test[84] were used to assess the risk of small-study bias for comparisons including an adequate number of studies.

Results

Search results

A total of 41 RCTs (27 from Cochrane reviews, 11 from other reviews and three additional trials from MEDLINE) met the selection criteria and were included in the evidence synthesis. Four RCTs investigated the clinical effectiveness of advice and information, two investigated simple analgesics, four investigated topical NSAIDs and 31 examined exercise interventions. A full list of references and summary of the main characteristics of the included RCTs are available from the authors of this report.

The knee was the most commonly affected joint among the RCTs considered in this review, with 24 RCTs investigating the knee only and 17 enrolling participants with either knee or hip OA. No trials investigating hand or foot OA met the inclusion criteria for this evidence synthesis and none assessed treatment effects independently of the joint affected.

Risk of bias

In general, the RCTs included in this review appeared to have good methodological quality and none of the medication trials (analgesics or NSAIDs) showed a high risk of bias on any of the domains. All 41 trials were judged to be at low risk of bias in terms of random allocation of interventions. Among four trials investigating advice and information interventions, one was at a high risk of bias in terms of loss to follow-up. For exercise interventions, the proportion of RCTs at a high risk of bias was 7 out of 31 (23%) for blinding of outcome assessment and 5 out of 31 (16%) for loss to follow-up.

Clinical effectiveness of primary care interventions for osteoarthritis

The 41 RCTs included in this review provided data on 6715 subjects assessed for pain and 5322 subjects assessed for functional limitation. *Table 2* summarises the results of the meta-analysis for each of the four comparisons.

Only one of the four RCTs investigating advice and information showed statistically significant improvement when compared with control. There was no evidence of heterogeneity across the studies for either pain ($I^2 = 0.0\%$) or functional limitation ($I^2 = 1.7\%$) outcomes. The pooled analysis showed a small, statistically significant, reduction in pain (SMD = –0.17, 95% CI –0.31 to –0.03, $n = 771$) as well as a small, statistically significant, improvement in function (SMD = –0.20, 95% CI –0.34 to –0.06, $n = 771$) in favour of advice and information (mean duration of treatment was 19 weeks) when compared with control.

The two RCTs investigating the effectiveness of analgesia (paracetamol, mean duration of treatment was 9 weeks) failed to demonstrate a significant difference compared with control interventions. The pooled

TABLE 2 Summary of the results of meta-analyses for the four main comparisons

Comparison	Pain outcomes				Functional limitation outcomes			
	Number of trials	Number of participants	I^2 (Cochran's Q-test for heterogeneity *p*-value)	Pain SMD (95% CI)	Number of trials	Number of participants	I^2 (Cochran's Q-test for heterogeneity *p*-value)	Functional limitation SMD (95% CI)
Advice and information vs. no treatment	4	771	0.0%, $p = 0.961$	–0.17 (–0.31 to –0.03)	4	771	1.7%, $p = 0.384$	–0.20 (0.34 to –0.06)
Analgesics (paracetamol) vs. advice/placebo/no treatment	2	400	–	–0.11 (–0.31 to 0.08)	1	57	–	–0.01 (–0.53 to 0.51)
Topical NSAIDs vs. advice/ placebo/no treatment	4	790	0.0%, $p = 0.985$	–0.35 (–0.49 to –0.21)	4	789	0.0%, $p = 0.900$	–0.31 (–0.45 to –0.17)
Exercise vs. analgesics/ advice/no treatment	31	1389	0.0%, $p = 0.990$	–0.32 (–0.43 to –0.21)	27	1240	0.0%, $p = 0.726$	–0.27 (–0.39 to –0.16)

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

effect estimates were small and not statistically significant: SMD = –0.11(95% CI –0.31 to 0.08, $n = 400$) for pain and –0.01 (95% CI –0.53 to 0.51, $n = 57$) for function. Given the small number of studies, these estimates have to be interpreted with caution.

Four RCTs compared the efficacy of topical diclofenac with placebo, with three RCTs showing beneficial effects. There was no evidence of heterogeneity of effects among the studies for either pain ($I^2 = 0.0\%$; $p = 0.985$) or functional limitation ($I^2 = 0.0\%$; $p = 0.900$). Pooled estimates of SMD showed moderate reduction in pain (–0.35, 95% CI –0.49 to –0.21, $n = 790$) and moderate improvement in function (–0.31, 95% CI –0.45 to –0.17, $n = 789$) of topical diclofenac compared with placebo after a mean duration of treatment of 5.5 weeks.

A wide range of exercise interventions was assessed among the 31 exercise trials, comprising both individual and group programmes, and with variable content focusing on strength, flexibility, balance and/or aerobic exercises. These references are included in *Figure 3*. Three RCTs showed a statistically significant reduction in pain (see *Figure 3*) and a statistically significant improvement in functional limitation. There was minimal statistical heterogeneity among the studies for both outcome measures (see *Table 2*). The pooled analyses showed a moderate reduction in pain (SMD –0.32, 95% CI –0.43 to –0.21, $n = 1389$) and moderate improvement in function (SMD –0.27, 95% CI –0.39 to –0.16, $n = 1240$) for exercise (mean duration of 19 weeks) compared with control interventions.

The number of trials of exercise interventions was sufficient to reliably assess the risk of small-study bias. The funnel plots for both outcome measures appear to be symmetrically shaped (*Figure 4*) and Egger's bias estimates confirm the lack of evidence for small-study bias for pain (bias estimate –0.49, 95% CI –1.41 to 0.44; $p = 0.292$) and function (bias estimate –1.14, 95% CI –2.47 to 0.18; $p = 0.087$).

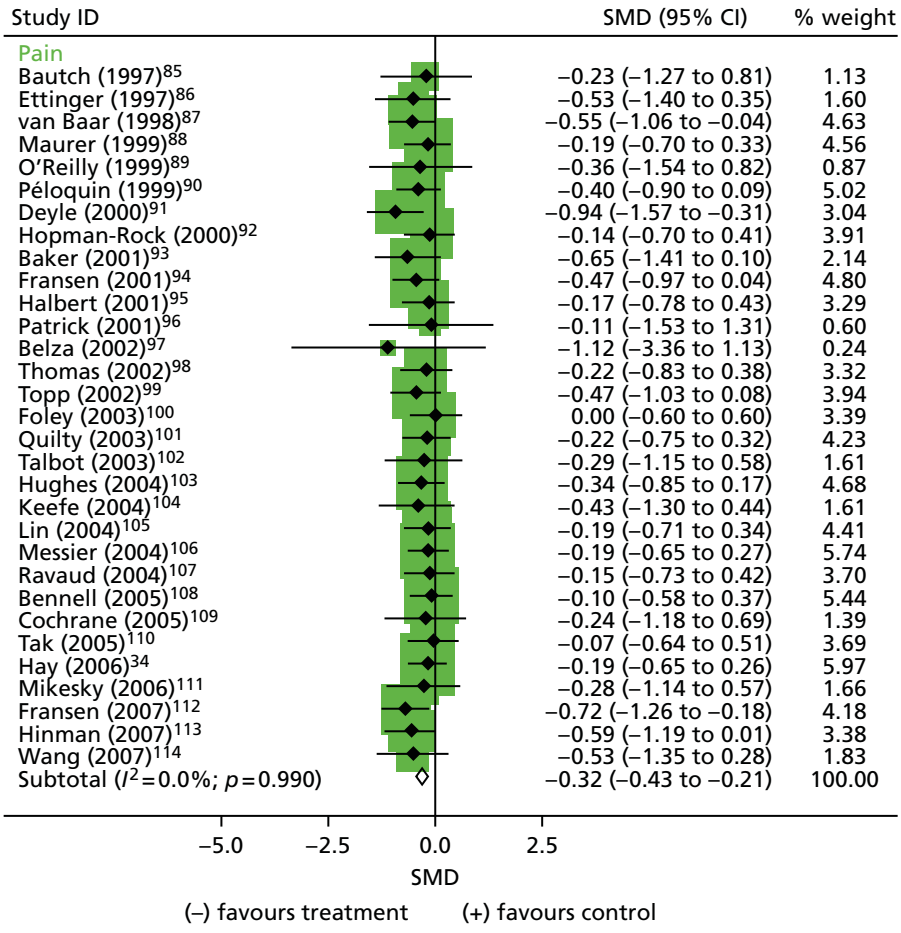

FIGURE 3 Effect estimates (SMD) for pain outcomes of exercise interventions for OA. Reproduced with permission from Jerome Wulff.[59]

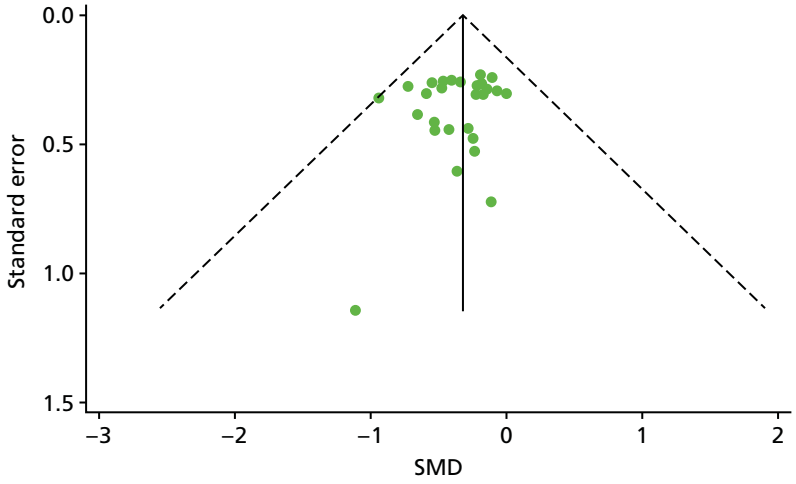

FIGURE 4 Funnel plot for trials investigating exercise interventions for OA (pain outcomes). Reproduced with permission from Jerome Wulff.[59]

Sensitivity analyses were performed for advice and information, and exercise intervention studies, with results showing minimal differences in the pooled effect estimates after exclusion of studies with a high risk of bias.

Discussion

This evidence synthesis and meta-analysis involved summarising the results of RCTs that had evaluated the effectiveness of advice and information, simple analgesics, topical NSAIDs and exercise interventions for primary care patients with OA. All trials investigated knee and/or hip OA. The meta-analyses demonstrated statistically significantly more reduction in pain and functional limitation for information and advice, topical NSAIDs and exercise interventions than control treatments. This was not the case for simple analgesics, namely paracetamol. The effect sizes varied (0.17 and 0.49) and may be rated as small to moderate.

Although our evidence synthesis only included trials carried out in primary care or open access settings, the results of the meta-analyses are in agreement with findings of previous or subsequent reviews on advice and information,[115] medication,[116–119] and exercise[120–122] in people with hip or knee OA, which show effect sizes for pain and function of similar magnitude.

This evidence synthesis showed that a minority of RCTs have been carried out in primary care settings, in particular those on advice and medication. Furthermore, this synthesis was mainly based on a search of available systematic reviews and only included full-trial reports published in English. This is a weakness of this review but it is reassuring that the magnitude of effect sizes were comparable with those reported in other systematic reviews based on broader search strategies, that is, including studies from health-care settings other than primary care and, therefore, not the focus of this review and evidence synthesis.

The quality of the RCTs included in the evidence synthesis appeared to be good, although information was often lacking regarding the methods used for concealing treatment allocation, blinding of outcome assessment and loss to follow-up. We accept that single reviewer assessment of quality was a weakness in the methods of this evidence synthesis, but sensitivity analyses excluding RCTs with a high risk of bias for at least one domain resulted in very similar effect estimates, which justifies use of estimates based on all available evidence.

Conclusion

This work provided effect estimates for the four primary care interventions included in the decision-modelling study (see *Study 3: estimating cost-effectiveness of delivering core primary care management for osteoarthritis*), which aimed to estimate the cost-effectiveness of different strategies to deliver core primary care management

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

for OA. The main shortcomings of the available evidence for this particular purpose concerned the relatively small number of trials carried out in primary care, the lack of trials focusing on hand or foot OA and the lack of long-term follow-up in available primary care trials.

Outputs

1. As stated above, estimates were provided for the modelling in the section *Study 3: estimating cost-effectiveness of delivering core primary care management for osteoarthritis*
2. The specific aim and content of the review presented here was extended into a full review and network meta-analysis of randomised trials of exercise interventions for OA as a linked output of workstreams and workstream 3.[122]

Study 3: estimating cost-effectiveness of delivering core primary care management for osteoarthritis

This section presents work undertaken to develop a decision model for optimal OA primary care and provides a basic template that can be developed and further refined using new or accumulating evidence on the cost-effectiveness of interventions.

The model aimed to evaluate the cost-effectiveness of two hypothetical interventions for delivering core primary care compared with current care for adults with OA, from a health-care perspective.

The two modes of delivery chosen to illustrate the template were (1) stepped care and (2) a 'one-stop shop', with both interventions containing four of the core primary care interventions for OA: advice and information, paracetamol, topical NSAIDs and exercise. The individual interventions were selected because they are recommended by NICE,[15] European League Against Rheumatism (EULAR) and Osteoarthritis Research Society International (OARSI)[10–14,32] as core interventions for managing OA. Therefore, the evidence synthesis described in study 2 (see *Study 2: summarising evidence regarding the effectiveness of core primary care interventions*) did not influence selection of interventions for the modelling study (paracetamol was found to be no more effective than placebo or alternatives, as judged by statistical significance, in the synthesis), but the synthesis did provide the standardised effect estimates for all four interventions included in the modelling study. The two hypothetical strategies were proposed by a consensus meeting consisting of clinicians and OA researchers.

Methods

Study population

The patient population was adults aged ≥ 50 years with symptomatic knee and/or hip pain or OA in a primary care setting in the UK, using data from the NorStOP cohort.[60] The original idea was to use a similar target population to that used for the prediction modelling study described above in study 1 (see *Study 1: predicting long-term outcome in people with joint pain and osteoarthritis*), in which participants had pain at baseline at one or more joint site (hand, hip, knee or foot). However, the proportion of participants with only hand or foot OA was low (< 5% each) and the evidence synthesis (described in study 2, see *Study 2: summarising evidence regarding the effectiveness of core primary care interventions*) provided effect estimates only for hip and/or knee OA. Therefore, it was decided to define this sample more narrowly than in study 1 of the workstream and the pain scores provided by the WOMAC questionnaire from study 1 were used to represent hip or knee pain status. If an individual had pain scores for both joints, the highest score was used to reflect the joint with greater pain.

Model structure

A Markov cohort model was built to reflect the clinical history of OA. The model considered four health states (no pain, mild, moderate and severe pain), defined using baseline WOMAC scores (range from 0 to 20).

A score of zero was defined as no pain, 1–5 as mild pain, 6–10 as moderate pain and 11–20 as severe pain.[62] Everyone started in the model with some pain (mild, moderate or severe) and moved health state when their condition improved or worsened; alternatively, they could remain in the same health state. Therefore, the model includes no pain as a state through which people could pass during follow-up, even though persons in that state were not included as one of the start points. Movement could only be to the next better or worse health state. A time horizon of 3 years was used in the base-case analysis, reflecting the follow-up period of the NorStOP cohorts used in this model. Death was not included in the model as the time horizon was short and the condition does not directly lead to death. A 3-month time cycle was used, as it was considered to be a clinically meaningful time period for OA in terms of expected changes in the symptoms of OA, duration of treatment and timing of decision-making by a GP. The structure is presented in *Figure 5*.

Interventions

Three packages of primary care were considered by the model: stepped care, 'one-stop shop' and current care. These hypothetical scenarios were discussed with clinicians to ensure assumptions were realistic.

Stepped care

This intervention included all four core interventions and assumed that interventions were prescribed by GPs in a stepped fashion. The first line of treatment was advice and paracetamol, and all patients were modelled as being offered this regardless of their baseline level of pain. If pain worsened or did not improve from a moderate or severe pain state, then the next (second) line of treatment prescribed was topical NSAIDs. The same principle applied for movement from the second to the third line of treatment (exercise, assumed to be supervised by a physiotherapist). If there was no improvement after the third line of treatment, patients returned to current care. Those who were originally in a severe pain health state and whose condition improved to moderate pain were moved to a 'moderate-from-severe' health state. If they remained in this 'moderate-from-severe' health state, the same line of stepped-care treatment was received, as maintaining any improvement from the most severe state was assumed to be a positive situation. Each step-up of treatment was assumed to involve a practice nurse appointment to introduce the new treatment.

'One-stop shop'

In the 'one-stop-shop' package of care, participants were offered all four core interventions simultaneously. It was assumed that the package was prescribed by a GP, with a physiotherapist offering the initial exercise package. The rationale behind this intervention was that it allowed patients to receive all optimal interventions at the same time, irrespective of pain severity. If a patient moved to a worse health state or did not improve from a moderate or severe pain state, they returned to current care. The rationale behind this assumption was that some participants will not respond or adhere to treatment and will then return to their GP to be placed under current care. The assumption regarding the continuation of an intervention once a moderate health state has been achieved after improvement from the more severe state also applied here.

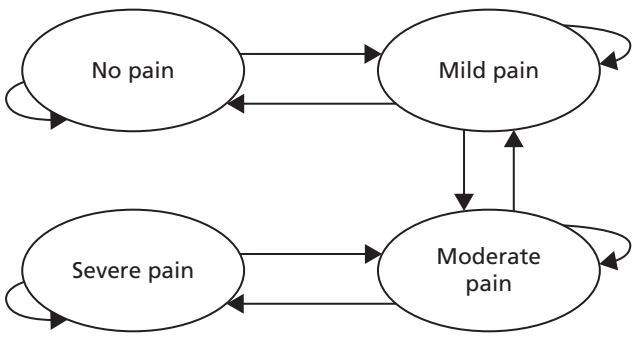

FIGURE 5 Markov model structure. Reproduced with permission from Jerome Wulff.[59]

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Current care

Current care for OA was informed by the observational cohort study (NorStOP) and a primary care-based RCT from the UK, which included a usual-care arm.[123,124] Data from the cohort study indicated that many patients only consult their GP about their condition once within a period of 3 years. Patients consulting their GP were likely to be initially offered advice and pain medication(s) according to the severity of their symptoms and, thereafter, stronger pain medications, referral for physiotherapy and eventually surgery as the last treatment option if the pain persisted. Surgery was not considered in this study given the 3-year time horizon, during which arthroplasty was not very likely. Owing to the lack of detailed data on health-care resource use in the cohort, the costs of current care were considered to be as reported in the control arm (current GP-led care) in a trial of exercise in knee OA patients.[123,124] Treatments included advice, exercise and pain medication(s) including simple analgesia (paracetamol and aspirin), topical and oral NSAIDs, and opioids.

Data inputs

Table 3 contains information on the baseline proportion of patients in each health state at baseline and provides estimates for the clinical effectiveness of individual treatments. The initial distributions in each health state at the start point of the model were estimated from NorStOP baseline data. Participants with no knee or hip pain at baseline were excluded because the model assumed only participants with pain would consult primary care and receive treatment. The numbers excluded from the sample used in study 1 above (i.e. those with only foot or hand pain) were small. Estimates of the SMD for each of the primary care interventions compared with their controls were obtained from the review in study 2.

Table 4 presents the transition probabilities for current care and all interventions. Current-care data on pain severity at baseline and 3 years from the NorStOP cohort was used to provide transition probabilities for the model. Matrix multiplication was utilised to transform actual transitions over 3 years into 3-monthly transitions. The treatment effect estimates were then applied to the current-care transition probabilities to obtain new transition probabilities. In the stepped-care intervention, only the effectiveness of the specific intervention at that step was applied, even though the patient was likely to be continuing to receive previous interventions. The estimate of effect (SMD) for the 'one-stop shop' was set at 0.5. This estimate was agreed during the consensus meeting of clinicians and OA researchers as it is larger than the strongest effect estimate of the individual primary care interventions. The clinicians in particular were clear that they often used the interventions as a package and agreed that a large additive effect between these concurrent interventions was unlikely.

TABLE 3 Model parameters

Parameter	Value	Source
Start point health state (pain)	***Baseline proportion***	
No pain	0	NorStOP data
Mild	0.3245	
Moderate	0.4177	
Severe	0.2578	
Treatment effect estimates	***SMD (95% CI)***	
Advice and information vs. no treatment	–0.17 (–0.31 to –0.03)	Systematic review (see *Table 2*)
Simple analgesia (paracetamol) vs. advice/placebo/no treatment	–0.11 (–0.31 to 0.08)	Assumption
Topical NSAIDs [diclofenac gel (Voltarol® gel, GlaxoSmithKline)] vs. advice/placebo/no treatment	–0.35 (–0.49 to –0.21)	
Exercise vs. analgesics/advice/no treatment	–0.32 (–0.43 to –0.21)	
'One-stop shop'	–0.50	

TABLE 4 Transition probabilities for pain health states

Pain health states by health-care category or intervention	No pain	Mild	Moderate	Severe
Baseline probabilities	0	0.3245	0.4177	0.2578
Current care				
No pain	0.6022	0.3978	0	0
Mild	0.0613	0.8836	0.0551	0
Moderate	0	0.0456	0.9194	0.0350
Severe	0	0	0.0462	0.9538
Advice and paracetamol				
No pain	0.6022	0.3978	0	0
Mild	0.1301	0.8148	0.0551	0
Moderate	0	0.0969	0.8681	0.0350
Severe	0	0	0.0980	0.9020
Topical NSAIDs				
No pain	0.6022	0.3978	0	0
Mild	0.2940	0.6509	0.0551	0
Moderate	0	0.2189	0.7461	0.0350
Severe	0	0	0.2215	0.7785
Exercise				
No pain	0.6022	0.3978	0	0
Mild	0.2729	0.6720	0.0551	0
Moderate	0	0.2032	0.7618	0.0350
Severe	0	0	0.2056	0.7944
'One-stop shop'				
No pain	0.6022	0.3978	0	0
Mild	0.4010	0.5439	0.0551	0
Moderate	0	0.2986	0.6664	0.0350
Severe	0	0	0.3021	0.6979

Apart from the baseline probabilities, each figure shown is the probability of transition from the column status to the relevant row status.

Costs

Table 5 shows the unit cost data applied in this study. The cost of current care for each pain severity health state was estimated using patient-level data from the current-care arm of the exercise trial by Hurley *et al.*[124] The unit cost of drugs was updated from a price year of 2003/4 to 2010 using *British National Formulary* (BNF) costs[126] and an average cost per 3 months of treatment calculated.

In stepped care, there was an initial cost for consultation with a GP, a nurse appointment for each new line of treatment and the cost of the intervention. The exercise intervention was six sessions of exercise at 30 minutes per session over a period of 6 weeks with an experienced physiotherapist.[128] It was assumed that when participants moved to the next line of treatment, they all continued to use the preceding interventions. The 'one-stop-shop' intervention cost included the cost of an initial GP consultation, ongoing

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 5 Costs (price year 2010)

Parameter	Cost/person/3 months (£)	Source
Current care		
No pain	9.60	Hurley *et al.*[124]
Mild pain	24.00	Hurley *et al.*[124]
Moderate pain	39.70	Hurley *et al.*[124]
Severe pain	65.70	Hurley *et al.*[124]
Intervention		
Advice in GP consultation	28.00	PSSRU[125]
Paracetamol	7.78	BNF[126]
Topical NSAIDs	16.17	BNF[126]
Exercise	34.75	Whitehurst *et al.*[127]
Staff costs		
GP consultation	28	PSSRU[125]
Nurse-led consultation	14	PSSRU[125]
Physiotherapist	34	PSSRU[125]

PSSRU, Personal Social Services Research Unit.

costs for paracetamol and topical NSAIDs, a consultation with an experienced physiotherapist to introduce the exercise intervention at an initial consultation and the cost of exercise. Costs were accumulated over the 3-year time horizon to give total costs for each of the three modelled treatment strategies.

Outcomes

Baseline Short Form questionnaire-6 Dimensions (SF-6D) component scores of the NorStOP participants with hip and/or knee pain were used to derive utility values for each health state, using the algorithms developed by Brazier and Roberts.[128] *Table 6* shows the mean utility scores for the four health states used in this study. Utilities were accumulated over the 3-year time horizon to give total quality-adjusted life-years (QALYs) for each strategy in the model.

Analysis

Base case

An incremental cost–utility analysis was undertaken to compare the cost-effectiveness of the two proposed primary care strategies with current care, from a health-care perspective. The interventions were ordered in descending order according to cost. Costs and QALYs were discounted at an annual rate of 3.5% in

TABLE 6 Baseline utility scores for health states[a]

Variable	Mean (95% CI)
No pain ($n = 131$)	0.7925 (0.7706 to 0.8144)
Mild pain ($n = 964$)	0.7604 (0.7518 to 0.7690)
Moderate pain ($n = 1193$)	0.6886 (0.6803 to 0.6969)
Severe pain ($n = 718$)	0.5593 (0.5499 to 0.5687)

a Source: NorStOP data at baseline.[60]

accordance with the current UK treasury guidelines.[129] Costs were expressed in Great British pounds using 2010/11 as the price year. A threshold of cost-effectiveness of £20,000 per QALY gained was adopted for this study.[15]

Deterministic sensitivity analyses

Sensitivity analyses were carried out to test the robustness of the primary results by changing some of the most important assumptions used in the model construction.

The following sensitivity analyses were performed:

- Extension of the time horizon of the model from 3 years to 5, 10 and 20 years.
- Application of GP costs instead of nurse costs for subsequent consultations in stepped care.
- Selection of only those subgroups with moderate or severe pain categories (excluding mild pain patients) to reflect a health-care seeking population.
- Varying the effect size (SMD) of exercise using 95% CIs.
- Varying the effect size (SMD) for 'one-stop shop' between –0.4 and –0.6.

Results

The results of the base-case analysis are presented in *Table 7*. Current care and the 'one-stop shop' intervention are both dominated by stepped care, which is cheaper and more effective. However the difference in QALYs is marginal. The 'one-stop-shop' intervention was cost-effective when compared with current care, with an incremental cost-effectiveness ratio (ICER) of £1341 per additional QALY gained.

Stepped care continued to dominate in all but one of the sensitivity analyses and the one-stop shop was always more cost-effective than current care. If a patient were assumed to visit a GP rather than a nurse at every change of treatment in stepped care, mean costs were only marginally increased (£24). Running the model for an equal proportion of patients with moderate or severe pain increased the costs of all the three interventions and decreased total QALYs. The lower effectiveness estimate for exercise slightly decreased the QALYs for stepped care. If the SMD for 'one-stop shop' is changed to –0.4 then stepped care still dominates, but if it is changed to –0.6 then total QALYs become greater than for stepped care, giving an ICER of £29,281 per QALY gained. This value is above the suggested lower threshold for NICE of £20,000/QALY and so 'one-stop shop' is still unlikely to be interpreted as cost-effective.

Discussion

This is the first model-based economic analysis of primary care interventions for knee and hip OA. The results of this study have demonstrated that the hypothetical stepped-care intervention was the most cost-effective, dominating 'one-stop-shop' and current care. This result was robust to the sensitivity analyses conducted.

TABLE 7 Base-case results

Treatment	Cost (£)	QALYs	Cost difference	QALY difference	ICER (£/QALY)
Comparison of all three options					
Stepped care	393.00	2.01	–	–	–
Current care	427.15	1.94	34.16	–0.06	(Dominated)
One-stop shop	507.62	2.00	114.62	–0.01	(Dominated)
One-stop shop vs. current care					
Current care	427.15	1.94	–	–	–
One-stop shop	507.62	2.00	80.47	0.06	1341.17

ICER, incremental cost-effectiveness ratio.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Inevitably, building a model of this type with hypothetical strategies has required a large number of assumptions to be made. First, the data used to provide transition probabilities for current care were available only for baseline and 3 years, and were converted to 3-month probabilities. It is highly unlikely that the long-term trajectory of 3 years can represent the short-term fluctuations in pain every 3 months. In addition, one-third of the cohort participants were lost to follow-up. As explained in *Characteristics of participants*, these participants had a slightly different age and sex distribution from responders. However, their other baseline measures (sociodemographic, pain, function, physical and mental health) were similar, limiting the likelihood of selection bias. The model was run for only 3 years in the base case, a short time horizon for a Markov model, although sensitivity analyses were conducted for long time periods. A longer time horizon would require the model structure to incorporate both joint replacement and all-cause death, both omitted from this model.

Second, assumptions were included regarding the effectiveness and cost of treatments. The level of effect of an intervention was assumed to be the same irrespective of the current pain health state, and the effect was assumed to be constant over time, even though estimates were from studies with short time frames. In reality, different types of patients may have different levels of benefit from interventions and the effectiveness of interventions may decrease over time. As patients moved from one treatment to the next in stepped care, there was no additive effect of treatment, even though it was assumed that the patient would still be receiving the previous intervention. Finally, the effect estimate for 'one-stop shop' was obtained from expert opinion to reflect the strength of the combined effect of the interventions. The value is likely to be reasonable, given that an additive effect is improbable when several interventions are combined. The current-care cost estimate used in this study was obtained from a data set for chronic knee pain, although this model considered both knee and hip OA.

Further assumptions were made with regard to patient pathways. When all treatments in the hypothetical interventions ceased to provide any improvement, the model assumed that patients moved back to current care in the next 3-month cycle, with a worse patient trajectory. In reality, this time frame may be too short. GPs may, for example, persevere with current interventions for longer and patients may be offered stronger medications, such as opioids, or be referred for surgery. The nature of the model does not allow the pain history of a patient to be taken into account. This limitation could be resolved by creating additional Markov health states or changing the type of model to one that can be run with individual patient histories (e.g. an Individual Sampling Model).[130]

Despite these many assumptions, the results concur with a common-sense view of how these alternative strategies for primary care interventions in OA would work out in practice. The basic difference between 'one-stop shop' and 'stepped-care approaches', based on the same set of possible core interventions, is that the former strategy will, by definition, mean that all consulters will receive all possible interventions, whereas the latter strategy can potentially be more parsimonious in restricting the proportion of people who will receive all four interventions.

Conclusion

Building this model and estimating the cost-effectiveness of two hypothetical primary care intervention strategies illustrates that economic modelling can be used to extrapolate beyond short-term clinical trials in OA. The results presented here are indicative of the type of output that can be produced and provide a template for future analysis. As they stand, they are not designed to provide immediate guidance for decision-makers.

However, it is likely that current care for some OA patients already follows either a stepped care or 'one-stop-shop' approach, although its content and delivery may differ from the assumptions made in the hypothetical models. Therefore, this work provides a platform for modelling long-term cost-effectiveness of interventions for OA, to which additional data can be added.

Outputs

1. The template developed, described and tested here is a stand-alone resource for future modelling exercises.
2. The practical application and content of the template described here was constrained by available data at this stage of the programme (population data and secondary care trials). A second application of the model will use the output of the primary care trial data, such as that which has now emerged from the later stages of the programme in workstreams 2–4 for new analyses in the future. This second application will also further develop the template through:

    i. probabilistic sensitivity analysis, taking into account all parameter uncertainty simultaneously, which was not undertaken here in this first application
    ii. inclusion of data on the course of OA pain over time, including short-term fluctuations, to provide more appropriate transition probabilities
    iii. adaptation to look at a broader range of OA pain sites.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Chapter 3 Workstream 2: The MOSAICS studies

Parts of this chapter have been reproduced from Dziedzic *et al.*[131] © 2014 Dziedzic *et al.*; licensee BioMed Central Ltd. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly credited. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated.

Abstract

Background: UK NICE guidance for primary care of OA patients is not implemented in practice.[15,24]

Aim: To develop and evaluate novel ways to deliver core NICE guidance for OA in primary care, and to investigate their implementation and clinical effectiveness and cost-effectiveness.

Methods: A suite of mixed-methods studies within the framework of a cluster RCT, in which four general practices implementing a new model of OA care (MOSAICS) were compared with four control practices.

Results: Pre randomisation: practical support for delivering NICE guidelines was developed, including training in model OA consultations for GPs and practice nurses, and OA patient guidebook developed with patients and health-care professionals.

A novel pop-up computerised template (the 'e-template') to record quality markers of care was implemented in the eight participating practices[132] and associated with improved recording of components of quality OA care and changed clinician prescribing and referral behaviour.

Post randomisation: the MOSAICS approach improved delivery of NICE core treatments, including provision of written information, simple analgesia and reduced radiography and increased physiotherapy referrals.

Only a minority of patients were referred to the practice nurse-led service. Among patients followed up post-consultation, MOSAICS innovations were regarded positively (by patients and health-care practitioners), pain and disability did not improve compared with controls, and visits to orthopaedic specialists and time off work declined.

Problems of implementation included difficulty maintaining change over time (patients and practitioners), practitioner variation in recording care and variable patient response to nurse-led clinics.

Conclusion: Improved implementation of NICE OA guidance can be achieved in primary care at no incremental cost, but does not result in better patient-reported pain and disability.

Ethics permission: This study was approved by the North West 1 Research Ethics Committee (REC), Cheshire (REC reference 10/H1017/76).

Trial registration: Current Controlled Trials ISRCTN06984617.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Introduction and overview of the MOSAICS studies

In the UK, OA is the second most frequent reason for consultations with older patients in primary care.[3] Most such consultations are with a doctor (the GP or family practitioner).[133] Primary care provides the arena in which most patients with OA in the population who seek care from the UK NHS are seen and managed.

In 2008, the UK NICE identified evidence-based interventions for patients with OA consulting in primary care (*Figure 6*).[15] These included a core set of interventions considered potentially applicable to all patients consulting about OA in primary care (represented by the inner circle in *Figure 6*).

Evidence has shown that most OA patients in primary care were not receiving or continuing with these interventions[24] and that there was widespread dissatisfaction and pessimism among doctors, health professionals and patients about primary care management of OA and about the potential to alter the natural history of the minority who will progress to joint replacement surgery.[22,41,42,44]

This context provided the ambition and purpose for this second workstream of the NIHR programme RP-PG-0407-10386. The workstream comprised a suite of studies (MOSAICS) designed to develop practical evidence-based ways to support the implementation of NICE guideline core interventions in primary care and, thereby, to enhance the value of the primary care consultation for patients with OA.

The focus of the NICE core guidance is on self-management.[15] The guidance characterises self-management in terms of access to information and advice, including specific advice about exercise and physical activity, optimal use of simple oral and topical analgesia and, when appropriate, weight loss. The OA research user group (RUG) at Keele (see *Appendix 1*) had highlighted, prior to submission of the NIHR programme proposal, that their main concern about NHS recommendations for self-management was, from the perspective of patients, the lack of information, support and help to adopt and maintain self-management approaches. This user view provided the underlying rationale for the content of the MOSAICS interventions. A model OA consultation for GPs in primary care backed up by new patient information, an electronic template for monitoring quality of care and referral for a series of practice nurse-led model OA consultations formed the 'MOSAICS model of care', and this was defined as the vehicle to deliver NICE guidelines into practice. The theoretical basis for developing the intervention to change GP behaviour is described in detail by Porcheret *et al.*[134]

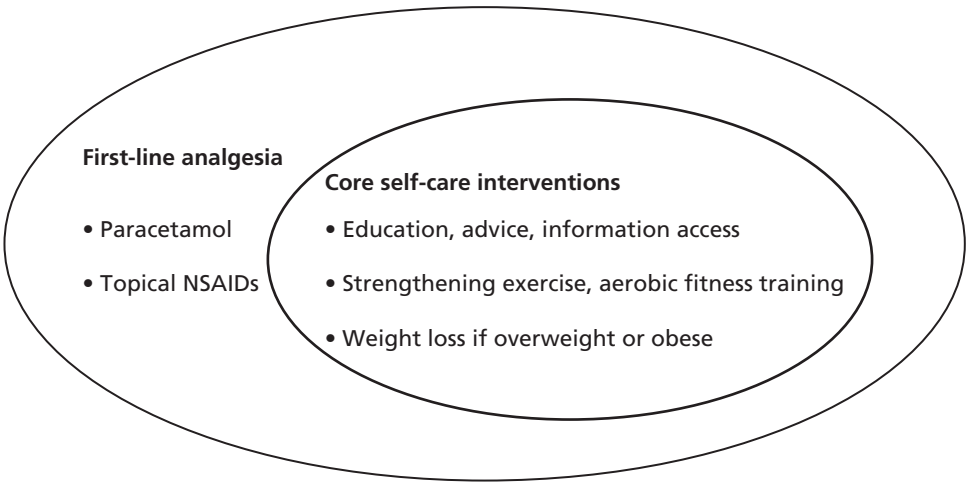

FIGURE 6 The NICE core recommendations for treatment of OA (redrawn using relevant components from NICE[15]).

The components of the MOSAICS approach were based on the Whole systems Implementing Self-management Engagement (WISE) model,[19] which embraces the needs of the patient, the health-care professional and the service. The key components in MOSAICS that were developed and evaluated in this workstream were as follows:

- Model OA consultations by GPs and practice nurses.
- Training for primary care health professionals to deliver model OA consultations.
- A new guidebook for patients, developed by patients and health-care professionals together, as the principal source of written information.[135]
- An electronic computer-based method (the 'e-template') for routine recording of components of quality OA care by GPs and practice nurses.
- New nurse-led OA clinics in general practice.

The suite of individual studies developing and evaluating different components of the MOSAICS model were embedded in a framework provided by a cluster RCT. The trial investigated implementation of the new model of care at a practice level (process outcomes) in four primary care practices that adopted the different components of the model OA consultation compared with four practices that (with the exception of the e-template) did not, as well as the impact of this implementation on pain and disability in a subgroup of patients recruited into a follow-up study (clinical outcomes). The trial was based on UK MRC complex trial principles.[136]

A second framework for studying aspects of adoption and implementation of the model OA consultation was provided by normalisation process theory (NPT).[137] Adoption of new or complex interventions in primary care is influenced by factors at three levels: the clinical encounter, the management and organisation of the practice, and the wider health service context. Recognition of these influences is increasing,[138] but there are few concrete examples that have investigated them. In studying the implementation of MOSAICS, we examined how GPs and practice nurses made sense of this complex intervention, what this meant for the way they worked and how this had to change, as well as which factors from the three levels were influencing this 'sense-making'.[139] The framework supplied by these concepts from NPT was used in two ways. First, it provided a structure for the research team in its approach to engaging with the practices by identifying factors facilitating implementation and overcoming barriers to delivering the MOSAICS intervention. Second, it provided a framework for analysing and evaluating the process of implementing the intervention.

Each component of the model OA consultation was studied separately, using a mixture of quantitative and qualitative techniques in order to establish its rationale, develop and evaluate its content and associated training, and test its acceptability with patients and professionals. The implementation of the components of the new model was investigated at a practice level in the trial through process measures of frequency of delivery and recording of care, and at an individual level through qualitative studies and surveys of patients and health-care professionals. The clinical effectiveness and cost-effectiveness of the programme was evaluated in the individual-level analysis of the subgroup that was followed up by self-report questionnaire in the cluster RCT.

Patient and public involvement (PPI) was present at every step of the process through a dedicated OA RUG (see *Appendix 1*).

The framework and component studies of MOSAICS

Each MOSAICS study or analysis is presented separately below, grouped into chapter sections: background (group 1), development and testing of interventions and quality-of-care measures including the e-template (groups 2 and 3), implementation of the new model of care (practice-level analysis), clinical effectiveness and cost-effectiveness (individual-level analysis) (group 4), patient and practitioner experience (group 5) and dissemination (group 6). Each study has aims, methods, results and published and practical outputs. The cluster RCT framework and populations are shown in *Figure 7*. The total practice population for the practice-level analysis is shown in blue and the subsample of the total practice population for the individual patient-level analysis is shown in green.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

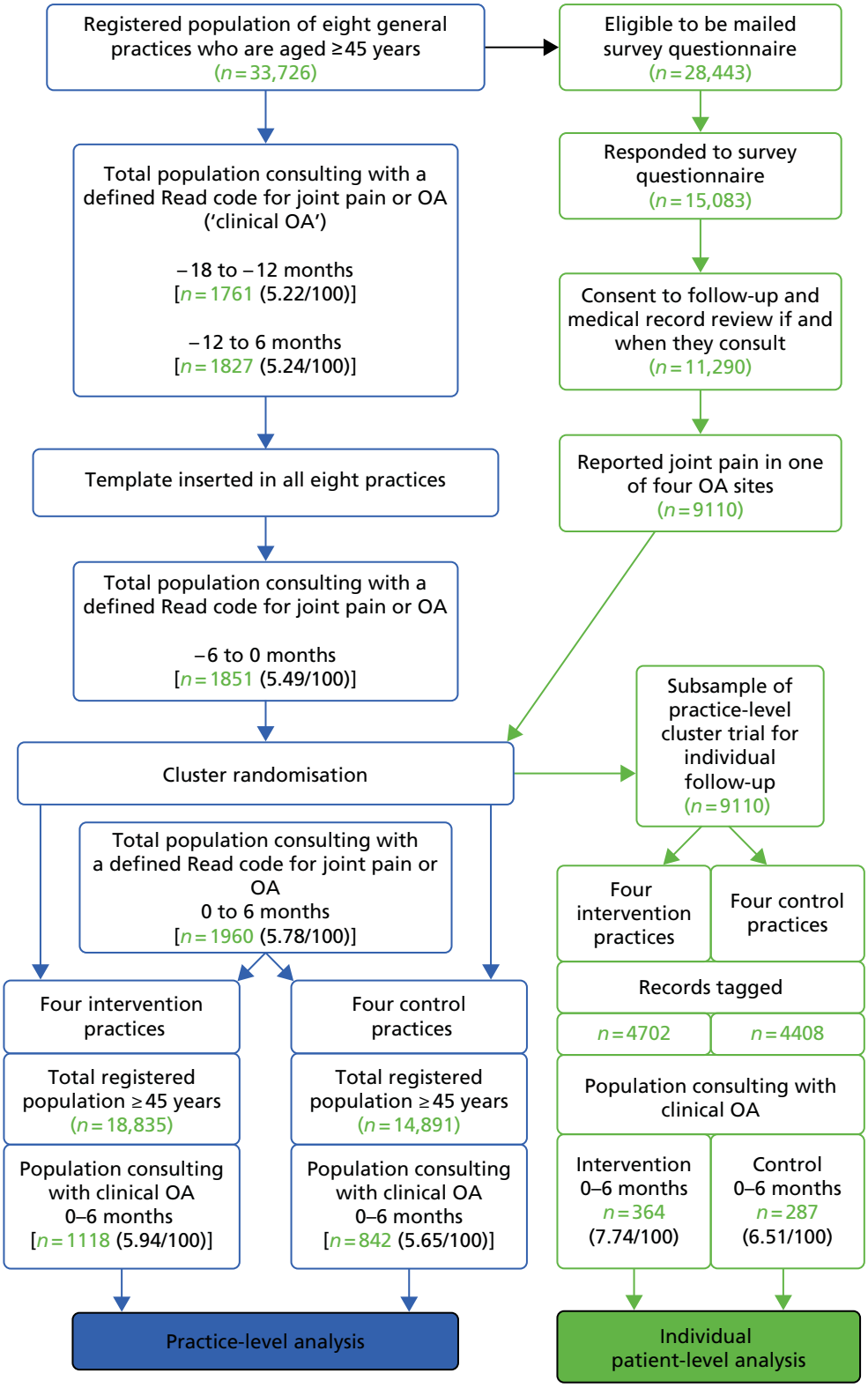

FIGURE 7 Cluster trial framework and population.

The study populations

The registered populations of the eight participating general practices who were aged ≥ 45 years ($n = 33,726$) provided the population base for the MOSAICS studies. The eligibility for the participation of these practices, their staff and their patients is listed in *Appendix 2* (see *MOSAICS eligibility criteria*).

There were three populations selected from this base in the pre-randomisation phase.

- Population survey responders. All persons aged ≥ 45 years registered with one of eight participating practices were sent a general health questionnaire to establish baseline pre-randomisation patterns of joint pain and OA in the population. A series of baseline analyses of all 'survey responders' were performed to determine the background frequency of multisite pain, self-reported uptake of NICE recommended treatments and prevalence of physical activity among persons with joint pain in this population (MOSAICS study 1.1, see *Study 1.1: self-reported prevalence and patterns of osteoarthritis and its treatment in the registered population of primary care*).
- Total OA consulting population. All consulters aged ≥ 45 years who were coded on the practice medical records system as having consulted about joint pain or OA, regardless of whether or not they had responded to the baseline population survey, formed the 'total OA consulter population' (*Figure 7* and *Box 2*). This anonymised population provided occurrence estimates of joint pain and OA in the participating practices, using the total registered population aged ≥ 45 years of the participating practices as the denominator ($n = 33,726$) (see *Figure 7*).
- Consenter population. Those who responded to the survey, who consented to follow-up and a recorded review, and who had reported pain in at least one OA joint site (hand, hip, knee or foot) at baseline were eligible for individual follow-up in the cluster trial. This population was a subgroup of the second population above, drawn from the baseline survey responders in the first population. A consent form to further contact and to use of individual medical records was included in the survey. Survey responders with pain in at least one joint who gave their consent formed the 'consenter population'. This identified, pre-randomisation, the pool of patients in all practices who would be willing to be recruited and followed up individually in the main cluster trial if they consulted their GP with joint pain during the randomisation phase. Such pre-randomisation of a consenting population of potential future eligible trial participants guards against bias in recruiting individuals within a cluster design post randomisation.[109] This subgroup, the self-reported questionnaires they completed during the follow-up phase and details from their medical records created the basis for the 'individual-level' analysis in the cluster RCT of clinical effectiveness and cost-effectiveness of the implementation of the new model of care (MOSAICS studies 4.2–4.3, see *Study 4.2: effect of the MOSAICS programme on patient-reported outcomes* and *Study 4.3: cost-effectiveness of the MOSAICS intervention: health economic evaluation*). The subgroup also provided the sampling framework for additional 'individual-level' quantitative and qualitative studies of implementation (MOSAICS studies 5.1–5.4, see *Study 5.1: implementing a complex intervention in practice – an evaluation using a theoretical framework*, *Study 5.2: experiences of model consultations – patients, Study 5.3: experiences of model consultations – general practitioners,* and *Study 5.4: experiences of the model consultation – nurses*).

BOX 2 'Read coding' of OA-related consultations

All UK GPs, including those in the participating practices, record on their practice computer software the reason why a patient consults them, using an established morbidity coding system called 'Read codes'. These codes can be symptoms or diagnoses. Preliminary work in MOSAICS by academic GP members of the team identified a complete list of symptom and diagnostic codes (summarised throughout this report as 'joint pain and OA codes') which were agreed as representing all possible presentations of clinical OA. The list of Read codes appears in www.keele.ac.uk/mrr,[140] Jordan *et al.*[141] and Edwards *et al.*[132]

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

A practice-level intervention (the quality-of-care e-template, which prompted OA assessment and NICE core OA interventions by the GP) was developed and introduced into all eight practices. This template was a pop-up computer screen that appeared whenever joint pain and OA consulting codes were entered. The effect of this intervention on routinely recorded items of care (such as prescriptions for OA) was investigated by a before-and-after comparison of downloaded anonymised practice record data 12 months before and 6 months after installation of the template in all eight practices [MOSAICS study 3.1, see *Study 3.1: the osteoarthritis quality-of-care template (the 'e-template')*]. In the subsequent cluster RCT of the MOSAICS intervention, this population (all patients consulting with OA symptoms or diagnosis in the participating practices) and the anonymised measures of quality of care provided by the template generated the level and method, respectively, of outcome analysis for the 'practice-level' implementation study. This compared anonymised downloaded consultation data between intervention and control populations (MOSAICS study 4.1, see *Study 4.1: implementation of the MOSAICS intervention – practice-level effect on quality of care for patients with osteoarthritis*).

Development studies

The development and testing of components of the model OA consultations and of the e-template were conducted in the pre-randomisation phase [MOSAICS studies 1.2, 2.1–2.5, 3.1 and 3.2, see *Study 1.2: establishing the current evidence base for multidisciplinary osteoarthritis interventions in primary care settings; Study 2.1: development of the content of a 'model' general practitioner consultation and the training intervention to deliver it; Study 2.2: evaluation of intervention workshops; Study 2.3: development of the content of a model osteoarthritis consultation with a practice nurse and the training package for nurses to deliver it; Study 2.4: evaluation of practice nurse training to support osteoarthritis self-management; Study 2.5: development of the opportunistic osteoarthritis consultation with health-care professionals; Study 3.1: the osteoarthritis quality-of-care template (the 'e-template');* and *Study 3.2: developing patient-reported outcomes with the osteoarthritis research user group*).

The main trial process

The eight practices were randomly allocated to intervention practices and control practices to form the cluster trial design. All practices continued to use the e-template to record care. Following training of GPs and practice nurses, consulters with joint pain or OA in the intervention practices received the MOSAICS model of care, whereas consulters with joint pain or OA in the control practices received usual care.

Practice-level outcomes

Practice-level outcomes in the cluster trial provided process measures of implementation of the model consultation in the intervention practices compared with controls and included:

- the proportion of consulters with joint pain or OA, whose e-template was completed in the post-randomisation period and who met the quality criterion for each item of care on the template
- the proportion who were referred to the nurse OA clinic
- changes in routine measures of OA care, such as prescribed medication and referrals to physiotherapy and radiography departments.

Individual-level outcomes

In the subgroup of individuals recruited within the cluster trial who were followed up post consultation at 0, 3, 6 and 12 months with self-report questionnaires, the extent of implementation of the MOSAICS model of care was further explored (e.g. in questions recalling the content of the consultations they had experienced, together with items on satisfaction with care). The main clinical outcomes of pain and disability were measured in this subgroup. Health economic analysis used both self-report questionnaires and data drawn from the medical records of this subgroup.

Qualitative interviews

Qualitative interviews were conducted with samples of patients after consultation and with health-care professionals delivering the new model of care (GPs and practice nurses) to explore experience and perspectives on the interventions.

Patient and public involvement in the MOSAICS studies

The OA RUG at Keele University had provided the initial impetus and thread for the whole programme (see *Appendix 1*) by highlighting that a desire expressed by patients to try self-management approaches was not matched by sufficient active and informed professional NHS support to do so.

The RUG continued to play a substantial part in the creation, shaping, delivering and evaluation of many important components of the MOSAICS suite of studies. An overview of RUG activity in MOSAICS is as follows.

Development and design of a guidebook for use in the osteoarthritis consultations

This was a stand-alone piece of work that pre-dated the NIHR programme RP-PG-0407-10386 and was published collaboratively by researchers and patients.[135,142] It was introduced and evaluated in the programme as a component of the model GP and nurse OA consultations. The OA guidebook differs from conventional patient education materials in that it contains lay, as well as biomedical, evidenced-based knowledge.

Five members of the OA RUG helped to shape the content of the guidebook. They reviewed a summary of qualitative research of people's experiences of living with OA to identify information needs, drawing on their own lived experience to suggest what information was required to meet the information needs of newly diagnosed patients. They also reviewed all draft materials.

Advice on content of the baseline population questionnaire

The aim of patients' involvement here was to critically assess a postal questionnaire that was used both as the basis for the background survey and as a means to recruit potential participants for the later follow-up stage of the cluster RCT. As a result of their feedback, changes to the questionnaire were made.

Involvement in the Delphi consensus study to inform the content of a model general practitioner consultation with osteoarthritis patients

The aim was to ask patients what they think should happen or should be done when older people with joint pain consult a GP for the first time. A 1-hour meeting was held to explain the consensus exercise and to discuss aspects of consultations. OA RUG members then completed two postal questionnaires that listed all the possible things that could happen during the consultation and decided which they thought should be included. For the first round, they were asked to consider this if time was no object. In the second round, they were asked what should be included in an initial 10-minute consultation. OA RUG members then returned for a follow-up meeting to discuss items on the questionnaires and to give feedback on the consensus method used to create the final content for the model consultations.[143,144]

Development of quality indicators for general practice consultations

A quality indicator describes the performance of something. It could be an action or task (e.g. GP recording information) or a heath outcome (e.g. lower blood pressure or reduced pain). Quality indicators are used to measure standards of care and whether or not they are consistent with what is thought to be best care. A list of quality indicators for the content of general practice consultations for the MOSAICS studies was defined by a systematic review.[145] This list was taken to the OA RUG and translated by the OA RUG into questions that could go into a questionnaire as part of the main study to ask patients about their experiences of the new consultation being tested. Therefore, some of the questions in this 'patient consultation questionnaire' are patient-defined questions (using the words suggested by OA RUG members).

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

This work by the OA RUG was also used as the basis to explore international comparisons of PPI methods. In collaboration with Norwegian, Danish and Portuguese colleagues, results of patient involvement across these three countries were compared. The format developed by the OA RUG was found to be almost identical to a format being tested for OA quality indicators in Norway, which also included patient partners.[146]

Analysis of data from the qualitative studies

Four members of the RUG assisted one of the qualitative research teams with the analysis of audio-recordings of nurse-led OA clinics. The RUG members were provided with transcripts of the clinics and asked to identify themes and to comment on the interactions between the nurse and the patients. The results of this were incorporated in the wider analysis of the clinics and were used to corroborate or amend the research team's existing coding frame for analysis.

Involvement in developing training for health-care professionals

At an OA RUG meeting in November 2011, researchers introduced members to four pieces of qualitative data from a previous study that investigated how people with knee pain self-managed their pain. These pieces of data were considered by researchers as good examples to use for training nurses in how people cope with joint pain and adopt self-management approaches. The OA RUG discussed and voiced their own thoughts on how patient experiences might influence a model OA consultation with the practice nurse and gave their critical consideration as to whether or not these examples would be suitable to help nurses to see the patient's view. The results of this session were fed into the training programme for the nurse-led consultation.

Four members of the OA RUG recorded extracts of patient stories for the training video for the practices.

Summary

The MOSAICS is a complex suite of studies. This created challenges for the RUG members and for researchers. We have managed this by:

- generating of a glossary which is available to all RUG members
- providing support for RUG members at meetings via a PPI co-ordinator and user support worker
- producing lay summaries in advance of meetings
- continuing to provide feedback to patients who have been involved
- offering training (e.g. Contributing Assertively in Meetings, a regular workshop run for staff by Keele University's Learning and Professional Development Centre).

The MOSAICS studies

The studies are presented separately below, grouped into background studies (group 1), studies of the development and testing of interventions and quality-of-care measures including the e-template (groups 2 and 3), implementation of the new model of care (practice-level analysis) and clinical effectiveness and cost-effectiveness (individual-level analysis) (group 4), patient and practitioner experience (group 5) and dissemination (group 6).

Group 1 studies: background evidence about osteoarthritis in primary care

Study 1.1: self-reported prevalence and patterns of osteoarthritis and its treatment in the registered population of primary care

This study incorporated three analyses of the baseline survey of the registered populations of the eight practices participating in the cluster trial. The survey was both a vehicle to identify a pool of consenters for potential future recruitment into the main trial if they consulted about joint pain or OA, and a basis for background studies of OA and its treatment in the general population.

The latter was the purpose of study 1.1 and is based on the fact that > 95% of people in the UK are registered with a general practice and so a survey of a registered practice population is a suitable vehicle for assessing a local general population sample.

Aims of the baseline survey

- To provide a well-characterised population sample of adults aged ≥ 45 years who consent to engagement in a follow-up study if they were to consult in the future about joint pain or OA.
- To provide the basis for three background analyses of OA, with the specific aims to:
    - describe the prevalence and characteristics of adults aged ≥ 45 years with multisite joint pain in the general population
    - determine the use of NICE recommended core OA treatments in a general population sample of adults aged ≥ 45 years with joint pain.
- estimate prevalence of physical activity in adults aged ≥ 45 years with joint pain in the general population.

Methods

Eligible adults on the register of one of eight practices identified for the cluster RCT were sent a letter of invitation to take part in the MOSAICS studies, together with information about the study and a copy of the baseline survey and a request to return the survey. On the last page of the survey there was a request for consent to further contact. Non-responders after 2 weeks were posted a reminder letter and survey. Those who did not respond or who indicated they did not wish to take further part were excluded from future engagement in the cluster RCT.

The questionnaire collected demographic and work-related data and information regarding general and psychological health, physical activity and joint pain in the 12 months prior to the survey as well as consultation behaviour and the management of their joint pain (hands, hips, knees and feet).

Figure 7 illustrates that the target population for the postal survey was 33,726 and, after exclusions by the practices, 28,443 were eventually mailed. There were 15,083 persons who completed and returned their questionnaire, a 53% response rate. Average age of responders was 63.9 years (SD 11.2 years) and approximately half were female.

Analysis 1: multisite peripheral joint pain – prevalence, impact and multidisciplinary support in community dwelling older adults

Parts of this text have been reproduced with permission from Andrew Finney.[147]

Introduction/background Patients present in primary care with symptoms rather than diseases and, if caused by musculoskeletal problems, such symptoms often take time to evolve into a recognisable form.[148] Therefore, people often live with joint pain without a formal diagnosis of OA. Joint pain is the most frequent reason for consulting with a GP in primary care.[149] Most publications focus on OA in single joint sites, such as the knee, overlooking the fact that most people with pain at one joint site will have pain in other joint sites concurrently.[54,150,151] This investigation used the baseline survey data to characterise patients with symptoms in multiple joint sites and aimed to provide context for the NIHR programme aim to develop optimal multidisciplinary primary care for patients with OA.

Objective To investigate the prevalence, distribution and impact of multisite peripheral joint pain in a community-dwelling population.

Methods 'Multisite' was defined as self-reported pain in two or more of the four main locations for OA pain in humans (hands, hips, knees, feet) using items enquiring about pain in the location during the past year.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Prevalence was defined as the proportion of people in the responder population who reported multisite pain. Distribution was investigated with respect to age, sex and social deprivation. Social deprivation was measured using linked census data that provides a small local area score (the Index of Multiple Deprivation) applied to the individual and is based on multiple characteristics of the locality area.[152] Multisite joint pain intensity was measured using a composite pain intensity score from four individual NRSs (scored 0–10) of the hand, hip, knee and foot, as well as total number of sites of joint pain reported and self-reported primary care consultations. 'Impact' of multisite joint pain was defined as cross-sectional association with scores from validated self-report instruments measuring physical [Short Form questionnaire-12 items (SF-12) PCS], mental (SF-12 mental health component score) and general health status as well as quality of life (QoL) [EuroQol-5 Dimensions (EQ-5D)].

Analysis Five subgroups of persons were formed, based on their reporting of the total number of joint locations with pain (i.e. 0–4). Descriptive statistics were used to describe the target population surveyed. Linear regression analysis was used to investigate the association of multisite peripheral joint pain with measures of general health, QoL and pain intensity, after adjusting for potential covariates [age, sex, body mass index (BMI) and social deprivation]. Logistic regression was used to calculate ORs as a measure of association between presence and intensity (moderate to severe) of multisite peripheral joint pain and recalled consultations with GPs, practice nurses and allied health professionals.

Results Among the 15,083 survey responders, 11,928 (79%) had any peripheral joint pain and 8206 (54%) had multisite peripheral joint pain. The distribution of multisite peripheral joint pain by number of locations involved is shown in *Table 8*, together with age and BMI.

Multisite pain was statistically significantly associated with increasing age, raised BMI and higher levels of social deprivation. It was also associated with increased pain intensity. Both the number of sites of peripheral joint pain and the levels of pain intensity were separately and independently associated with reduced general health and QoL (data not shown).

Self-report of primary care consultations for peripheral joint pain revealed that GPs were the main contact for this problem and that there was substantially lower reported contact with the wider multidisciplinary team. The odds of consulting with primary care health professionals were increased in those with multisite peripheral joint pain and in those with moderate to severe pain intensity. The likelihood of consulting a GP or a practice nurse with peripheral joint pain increased with each increase in the number of painful joint sites and was at its highest in those with pain in four sites (GP consultation: OR 4.51, 95% CI 3.98 to 5.11; practice nurse consultations: OR 4.30, 95% CI 3.41 to 5.39) compared with one site (the reference group for analysis, with OR assigned a value of 1.0). The odds of having consulted a GP or practice nurse with multisite peripheral joint pain rather than single site peripheral joint pain increased with age.

TABLE 8 Distribution of number of joint pain sites and mean (SD) age and BMI of persons in different categories defined by number of painful sites

Population characteristic	0 painful sites	1 painful sites	2 painful sites	3 painful sites	4 painful sites	ANOVA
Number of people (%)	3155 (21)	3722 (25)	3565 (23)	2688 (18)	1953 (12.9)	
Age (years)	62.6 (11.3)	63.0 (11.1)	63.8 (11.0)	64.8 (11.2)	66.3 (10.9)	$F = 43.6$**
BMI (kg/m^2)	25.6 (4.0)	26.3 (4.1)	27.0 (4.5)	27.7 (5.1)	28.4 (5.6)	$F = 138.1$**

**$p < 0.001$.
ANOVA, analysis of variance.

Conclusion Multisite peripheral joint pain is the most common form of joint pain in a UK general population sample of adults aged ≥ 45 years. In this study, it was associated with increasing age, raised BMI, higher levels of social deprivation, worse levels of general health, worse levels of QoL and higher pain intensity. The important result that is of relevance to the aim of the NIHR programme as a whole is that people consulting with the clinical syndrome of OA in primary care (i.e. peripheral joint pain affecting hip, knee, foot or ankle) are likely to have pain in multiple sites and yet their reported contact with all members of the multidisciplinary team who might provide treatments that get beyond a single site was very low.

Implications for optimal primary care in practice are as follows.

1. Training and education of primary care practitioners should highlight the need to establish the extent of joint involvement in patients with clinical OA, even when the presentation concerns a single joint.
2. Research, training and education should develop and evaluate methods of self-management relevant to pain and disability at multiple joint sites. This is likely to include multidisciplinary care, such as podiatry.
3. The effect of multidisciplinary care on multisite joint pain outcomes is a gap in the evidence.

Analysis 2: uptake of National Institute for Health and Care Excellence osteoarthritis core treatments in community-dwelling older adults who have consulted for joint pain

Parts of this text have been reproduced with permission from Healey EL, Afolabi EK, Lewis M, Edwards JJ, Jordan KP, Finney A, Jinks C, *et al.* Keele University. 2017.

Introduction/background The uptake of NICE OA core treatments[15] in community-dwelling older adults with a self-reported primary care consultation for joint pain is unknown.

Aim The aims of this study were to use the population survey to:

1. describe the uptake of the core NICE OA treatments[153] in a community-dwelling older adult population with a self-reported primary care consultation for joint pain
2. determine whether or not a self-reported clinical diagnosis of OA is associated with the uptake of these guidelines.

Methods Participants in the population survey provided information on the presence of joint pain (hand, hip, knee, foot) and also reported on consultations about (and management of) their joint pain over the last 12 months. Uptake of NICE recommended treatments in responders who consented to medical record review and reported a consultation for joint pain in the previous 12 months was compared between those with and without a coded diagnosis of OA in their medical records.

Results *Figure 7* shows that 9110 survey responders consented to medical record review and had reported joint pain in at least one of the target sites in the population survey. Of these, 4059 (44.6%) reported consulting primary care in the last 12 months and 502 of these (12.4%) had been given an OA diagnosis, rather than joint pain symptom codes only or no diagnostic code relevant to OA.

Reported uptake of the core non-pharmacological treatments was low. Responders reporting joint pain were more likely to report having received pharmacological treatments such as paracetamol and topical NSAIDs. Those with an OA diagnosis in their medical records were more likely to report having received treatment in line with the guidelines (*Table 9*). Responders aged ≥ 65 years were statistically significantly more likely to report the use of paracetamol than responders aged < 65 years ($p = 0.01$) and were less likely to report receiving the core non-pharmacological treatments of education and access to information ($p = 0.01$) and muscle-strengthening exercise ($p < 0.001$).

Conclusions Reported uptake of core non-pharmacological treatments for OA was lower than for recommended pharmacological treatments in this general population sample of those aged ≥ 45 years who responded to a postal questionnaire, reported having consulted in the previous year about joint pain

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 9 Reported use of NICE core treatments in consulters with and without an OA diagnostic Read code in their medical records

Treatment	Joint pain group, *n* (%) (*N* = 3557)	OA code group, *n* (%) (*N* = 502)
Written information about: treatment,* self-management,* OA***	358 (10.1), 361 (10.1), 208 (5.8)	68 (13.5), 70 (13.9), 61 (12.2)
Muscle-strengthening exercises, general aerobic fitness, weight loss[a]	452 (12.7), 130 (3.7), 230 (9.5)	80 (15.9), 24 (4.8), 31 (8.9)
Paracetamol,*** topical NSAIDs***	1067 (30.0), 889 (25.0)	203 (40.4), 166 (33.1)

*$p < 0.05$; **$p < 0.01$ ***$p < 0.001$.
a Recommended for overweight and obese individuals.

and gave consent to use of their medical records for the research project. Those with a diagnosis of OA in their medical records reported receiving treatment more in line with NICE recommendations than those without such a diagnosis, while older people with joint pain received proportionally more pharmacological than non-pharmacological treatments than younger people. This is in contrast to NICE guidance[15] recommending these treatments irrespective of age.

Implications for optimal primary care of patients with OA in practice are as follows:

1. Persons with the highest prevalence of OA and the highest risk of side-effects from pharmaceuticals [i.e. older persons (those aged ≥ 65 years)] have the lowest uptake of non-pharmacological OA therapies and, therefore, represent a specific target for development of ways to support self-management.
2. Research needs to explore the specific barriers or attitudes among patients and professionals towards adopting non-pharmacological approaches to OA management.

Analysis 3: self-reported physical activity levels – measurement and assessment in community-dwelling older adults with, or at risk of, osteoarthritis

Parts of this text have been reproduced with permission from Robert Smith.[154]

Introduction/background NICE recommends that physical activity should be a core treatment for all adults with OA.[15] Adults aged ≥ 45 years with symptomatic joint pain either are likely to have already developed OA or be at risk of OA. Joint pain is one of the main symptoms experienced by individuals with OA and yet in the UK the levels of physical activity in adults with joint pain or diagnosed OA are unknown.

Aim The aim was to estimate current levels of physical activity in adults with joint pain.

Methods This was an analysis of the baseline survey described above. A questionnaire item from the survey used for this analysis was the measurement of physical activity using the Short Telephone Activity Rating (STAR) questionnaire. The STAR questionnaire is scored corresponding to Department of Health (DH) guidelines for weekly physical activity levels[155] (participating in moderate physical activity for 30 minutes at least five times per week or participating in vigorous physical activity at least three times a week). Participants are scored as 'inactive' if they report participating in moderate or vigorous activity less than once a week, 'somewhat active' if they report participating in moderate or vigorous activity more than once a week but less than recommended by the DH, or 'active' if they report participating in moderate or vigorous activity that meets or exceeds DH recommendations.

This analysis included all participants who returned a baseline survey and who had fully completed the STAR questionnaire ($n = 11{,}310$). The prevalence of activity and inactivity in relation to the presence of joint pain and other characteristics was calculated, with associations estimated by logistic regression.

Results More adults with self-reported joint pain were inactive than those with no reported joint pain and fewer adults with self-reported joint pain achieved the recommended levels of physical activity than those with no reported joint pain. Participants with self-reported joint pain were less likely to be physically active than participants with no reported joint pain (OR 0.75, 95% CI 0.68 to 0.77).

Conclusion There is convincing evidence that physical activity is beneficial for older persons with joint pain and OA. However, in this population survey, current levels of physical activity were lower in older adults with joint pain than those without. This is a cross-sectional analysis and joint pain might have prevented physical activity or physical activity might have prevented joint pain. However, the results suggest that there is scope for substantial increase in activity levels in persons with joint pain and OA.

Implications for optimal primary care of patients with OA in practice are as follows:

1. There may be specific barriers to exercise as a means of self-management of OA.
2. Research needs to explore these barriers to exercise (e.g. patient fears of joint damage), the individual benefits of exercise and the potential and/or willingness for such barriers to be overcome.

Study 1.2: establishing the current evidence base for multidisciplinary osteoarthritis interventions in primary care settings

Parts of this section have been reproduced from Finney *et al.*[156] ©The Author(s). 2016. This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated.

A systematic review of randomised controlled trials using a multidisciplinary approach to managing the patient with osteoarthritis in multiple joint sites

Introduction/background The NICE OA guidance identified core treatments and highlighted the need to consider combination therapies for persons with OA in multiple sites.[15] The term 'combination therapy' suggests packages of care that cover a range of interventions delivered by a wide range of health disciplines (i.e. multidisciplinary care). There is as yet no evidence about the effects of a multidisciplinary package of care targeting multiple sites of OA.

Aim To undertake a systematic review to provide a synthesis of evidence for the clinical effectiveness and cost-effectiveness of NICE core treatments delivered by a multidisciplinary team (defined as two or more of the following: GP, practice nurse, physiotherapist, occupational therapist, podiatrist, community pharmacist, nutritionist, orthopaedic specialist) in primary care settings, targeting multiple sites of OA across the peripheral joints (hands, hips, knees or feet).

Methods Computerised bibliographic databases were searched from database inception until June 2013 [MEDLINE, EMBASE, Cumulative Index to Nursing and Allied Health Literature (CINAHL), PsycINFO, British Nursing Index, Health Business Elite, Health Management Information Consortium (HMIC), Allied and Complementary Medicine Database (AMED), Web of Science and The Cochrane Library]. Studies were included if they met the following criteria: a RCT, a primary care population with OA in at least two different joint sites, and interventions undertaken by at least two different health disciplines. The Cochrane Risk of Bias Tool was utilised for quality assessment.[82] A narrative analysis was used to summarise the findings.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Results The searches identified 1148 titles. The screening of titles reduced the number of papers down to 211. The review of abstracts left 79 eligible papers. A deductive process left five papers describing a total of three studies that were eligible for the review (including one protocol and one preliminary results paper). Reference checks of current guidelines found no studies beyond those found within the searches. Reference checks of a systematic review identified one further study that was eligible for the review. Therefore, the total number of eligible studies was four.

The four trials each used core interventions endorsed by OA guidelines. The trials targeted multiple sites of OA in either hips and knees, or hips, knees and hands, yet were very heterogeneous. None of the trials reported a multidisciplinary package of care deemed clinically effective or cost-effective for adoption in UK primary care, yet there were several moderately effective outcomes across the four trials. Educational approaches demonstrated moderate reductions in health service utilisation in primary care. Exercise interventions, although initially effective, were shown to decline over time. An interdisciplinary GP and practice nurse intervention showed the most positive outcomes.

Conclusion This systematic review identified four trials investigating differing multidisciplinary approaches for multiple-site OA. The review identified no strong evidence to support the adoption of any one of the trial methods as a package of care, but did identify positive aspects of each trial. The review highlighted that a consistent approach to outcome measurement is required for future studies of this nature as there is limited consensus on outcome measures at present, which leads to greater heterogeneity across studies.

Implications for optimal primary care of OA in practice are as follows:

1. This review gives some support for

    i. the effectiveness of non-pharmacological interventions for OA (education and exercise) in managing whole-person multisite OA
    ii. shaping the primary care OA team around nurses and GPs.

2. The review did not identify clear evidence for the effectiveness of the models of delivery that are already in place in practice. This provided an additional rationale for investigating new models of primary care in MOSAICS.

Group 2 studies: developing interventions to support self-management of osteoarthritis in primary care and to monitor the quality of care delivered

Study 2.1: development of the content of a 'model' general practitioner consultation and the training intervention to deliver it

Parts of this section have been reproduced from Porcheret *et al.*[134] © 2014 Porcheret *et al.*; licensee BioMed Central Ltd. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly credited. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated; and from Porcheret *et al.*[143] © 2013 Porcheret *et al.*; licensee BioMed Central Ltd. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Introduction/background

The MOSAICS studies required design and implementation of an intervention to enhance management for persons with OA in primary care. The MOSAICS OA model of care, which centred on consultations by GPs and practice nurses, was developed as a patient-focused complex intervention and designed for the assessment and treatment of older adults (those aged ≥ 45 years) presenting in general practice with

peripheral joint problems and to enhance support for self-management for people with OA in primary care. The content aligns with recommendations in the NICE OA guidelines.[15] It was implemented in the intervention practices of the MOSAICS studies.

The implementation necessitated a change in GPs clinical practice during consultations with patients with OA. The use of theory, systematically selected and used in the development and implementation of complex interventions, is recommended in the MRC guidance on complex interventions.[136,157] Implementation science has developed and tested theoretically derived models and frameworks for the development of behaviour change interventions to change clinical behaviour, and these were used as a basis for this first substudy.[134]

Aim

The aims of this substudy were to (1) develop the content of the model GP consultation for patients with OA and (2) use theory to develop the training package to enhance the delivery of the model OA consultation by GPs in the MOSAICS studies intervention practices.

Methods

Phase 1: development of the model OA consultation with a GP A postal Delphi consensus exercise was undertaken with two expert groups: (1) 15 GPs with expertise in OA management and (2) 14 patients with experience of living with OA. An advisory group generated 61 possible consultation tasks for consideration in the consensus exercise. The expert groups were asked to consider which tasks should be included in the model OA consultation. The level of agreement for inclusion in the model was set at 90%.

Phase 2: development of the GP training package The training package was developed as a behaviour change intervention to implement the model OA consultation and its development was guided by four theoretical models/frameworks: (1) an implementation of change model to guide the overall approach, (2) the Theoretical Domains Framework (TDF) to identify relevant determinants of change, (3) a model for the selection of behaviour change techniques to address identified determinants of behaviour change and (4) the principles of adult learning.

The Phase 2 development was undertaken in three steps.

Step 1: development of a concrete proposal for change The behaviour change required of the GPs was the delivery of the model OA consultation. Following phase 1 development of the model for the OA consultation, two activities were undertaken. First, the characteristics of the model OA consultation were compared with characteristics known to promote or hinder the implementation of an innovation. Second, three general practice advisory groups were formed and meetings arranged. The meetings were audiotaped and field notes made. The model OA consultation for GPs was presented to the groups and their views and understanding obtained. From the results of the comparison and feedback from the advisory groups, the model consultation was refined to enhance uptake by GPs.

Step 2: analysis of performance, target group and setting At the same meetings as those arranged for step 1, the advisory groups were asked about (1) their current management of OA, (2) their awareness of, and agreement with, the NICE OA guidelines[15] and (3) any gaps that they perceived between their current practice and that recommended by NICE and in the model consultation. In addition, they were asked to suggest which barriers and/or incentives might be relevant to implementing the model consultation in practice. Their responses were mapped by the study team to the domains in the TDF.

Step 3: development or selection of strategies and measures to change practice There were four parts to the development of the behaviour change intervention: (1) defining content, (2) selecting behaviour change techniques, (3) deciding on style of delivery and (4) addressing local practicalities. The content was developed by the study team and informed by the views of GPs from step 2. The mapping of

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

behaviour change techniques to TDF domains was utilised to select the techniques to address the domains identified in step 2. The adult learning principles and the Cochrane Effective Practice and Organisation of Care Group's reviews were used to decide on the style of delivery.[158] Practical issues, such as venues, timings and duration of meetings, how best to deliver the behaviour change intervention, and what was feasible in the MOSAICS studies, were addressed by the study team in consultation with the general practices in the study. More details of methods appear in Porcheret *et al.*[134]

Results

Phase 1 The model OA consultation included 25 tasks to be undertaken during the initial consultation between a GP and a patient presenting with peripheral joint pain. The 25 tasks provide detailed advice on how the following elements of the consultation should be addressed: (1) assessment of chronic joint pain, (2) patients' ideas and concerns about their condition, (3) exclusion of red flags, (4) examination, (5) provision of the diagnosis and written information, (6) promotion of exercise and weight loss, (7) initial pain management and (8) arranging a follow-up appointment. Both GP and patient groups prioritised a biomedical approach to the consultation, rather than a biopsychosocial model, suggesting discordance between current thinking and research evidence. The 25 tasks are detailed in Porcheret *et al.*[143]

Phase 2 The behaviour change intervention presented the GPs with a clearly defined proposal for change, it addressed seven of the TDF domains, it incorporated 10 behaviour change techniques and it was designed to be delivered in workshops that incorporated the expertise and professional values of the GPs. The programme for the workshops used a mixture of interactive and didactic sessions facilitated by opinion leaders, and utilised 'context-bound communication skills training'.

Conclusions

Phase 1 enabled expert, patient and GP views to be used to reach consensus on a model OA consultation. Phase 2 enabled the systematic and theory-driven development of the training package for GPs, which was designed as a behaviour change intervention to enhance their care for people with OA.

Study 2.2: evaluation of intervention workshops

Parts of this text have been reproduced with permission from Porcheret *et al.*[159] This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated.

Introduction/background

Evaluation of intervention workshops to change clinical practice was undertaken at three levels: (1) learner reaction to workshops, (2) clinical practice in a simulated setting and (3) clinical day-to-day practice.

Aim

The aim of this study was to develop methods and measures to evaluate the workshops developed in study 2.1 and to undertake the evaluation.

Methods

Methods and measures were developed for the three levels of evaluation.

Learner reaction to workshops A learner reaction questionnaire was developed and administered at the end of the final workshop. It encompassed level 1 Kirkpatrick educational outcomes, such as level of enjoyment, views on content and confidence in delivering the model OA consultation.

Clinical practice in a simulated setting GPs were video-recorded consulting with simulated OA patients at three time points: before the workshops, 1 month after the workshops and 5 months after the workshops. Detailed scenarios and biographies were developed for the simulated patients and training was undertaken so that the 'patients' presented to the GPs and responded to them in a consistent manner. Videos were rated using a specially developed instrument by four trained assessors whose accuracy had been established. Assessors viewed, in random order, all videos of their allocated GPs. The presence or absence of 14 consultation tasks was assessed for each GP at each time point. The 14 tasks used in the experiment are shown in *Appendix 2* (see *MOSAICS model consultation task list*).

Two measures were developed to evaluate change in clinical practice: a GP competency score which measured the extent to which a GP delivered the consultation (how many of the 14 behaviours were present in an individual video), and a task delivery score which measured the overall extent to which a task was delivered by the GPs (how many of the GPs undertook the task at a given time point).

1. Development of the rating instrument: draft rating criteria were developed based on key elements of the model OA consultation, which could be observed on a video-recording. The criteria were refined by a panel of four reviewers (academic GP, pain specialist, physiotherapist and epidemiologist) using two demonstration videos of the model OA consultation that were used in the assessor training sessions. The rating instrument contained criteria for rating 14 consultation tasks necessary for the delivery of the model OA consultation.
2. Establishing acceptability and accuracy of use of the rating instrument: gold standard ratings – behaviour present or absent – were produced by the panel for five video consultations of participating GPs obtained before and after the training. Four other GPs, who were not members of the practices participating in the training, learned to rate video-recorded consultations using the rating instrument. They then individually rated the five videos, which were presented to them in random order. For each rater, percentage agreement and sensitivity and specificity compared with the gold standard ratings were determined. Assessors were able to rate all the tasks for all five videos. The percentage agreement with the gold standard for the four assessors ranged from 80% to 86%. The sensitivity (proportion of tasks rated present if gold standard rated present) ranged from 85% to 98% and specificity (proportion rated absent if gold standard absent) ranged from 46% to 83%. The rank order of the videos by the number of tasks present was comparable between assessors and gold standard (*Table 10*). Further training was undertaken with the assessors to improve specificity.

Clinical practice in day-to-day primary care During the MOSAICS studies, an audit was undertaken of the extent to which four consultation tasks had been undertaken in daily practice by GPs in the participating practices. Patients who visited the GP for OA-related problems in the intervention practices,

TABLE 10 Assessor ranking of videos, ordered by gold standard rating rank

Video	Gold standard rating rank	Assessor 1	Assessor 2	Assessor 3	Assessor 4
D	1	1 =	1	1 =	1
C	2	1 =	3 =	1 =	3 =
A	3	3	3 =	1 =	2
B	4	5	3 =	4 =	3 =
E	5	4	5	4 =	5

=, of equal rank. The statistical analysis here uses rank order as the basis for testing.

Parts of this text have been reproduced with permission from Porcheret *et al.*[159] This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

and subsequently attended the nurse-led OA clinic, were asked to report on the content of the previous GP consultation. Four consultation tasks were asked, namely, did the GP (1) elicit ideas about the problem, (2) give the diagnosis, (3) explain the diagnosis and (4) hand out the guidebook?

Results

Learner reaction An extract of the results is shown in *Table 11*.

Clinical practice in a simulated setting Fifteen GPs were video-recorded consulting with simulated OA patients at the three time points. Median GP competency score at baseline was 7, at 1 month after the workshops it was 11 and at 5 months after that it remained at 11 (*Table 12*). Task delivery score increased for eight tasks after workshops; for other tasks the delivery score was either high before the workshops or not affected by workshops.

Clinical practice in day-to-day primary care This was undertaken as part of the main MOSAICS trial through patient questionnaires completed after their consultation with the GP about joint pain and is reported in *Study 2.4: evaluation of practice nurse training to support osteoarthritis self-management*.

Conclusions

Methods and measures were developed for the five levels of training evaluation. Learner reaction to workshops was positive and the workshops were associated with changes in GP consulting behaviour with simulated patients.

TABLE 11 Participant ratings of the GP consultation training and its delivery

	Number (%) of participants ($n = 23$)			
Statement	Strongly disagree	Disagree	Agree	Strongly agree
I enjoyed the training sessions			16 (70)	7 (30)
The training has helped me to better manage OA			14 (61)	9 (39)
The training covered a lot of ground I already knew		4 (17)	16 (70)	3 (13)
The training has helped with other aspects of my practice		6 (26)	13 (57)	4 (17)
The trainers were proficient in delivering the sessions			14 (61)	9 (39)
I would recommend these training sessions to a colleague		1 (4)	15 (65)	7 (30)

TABLE 12 Summary statistics for GP Competency Score by time point

Summary statistics	Baseline	1 month after workshops	5 months after workshops
Median	7	11	11
IQR	5–8.5	10–12	10–11
Range	5–11	8–14	7–13

IQR, interquartile range.

Parts of this text have been reproduced with permission from Porcheret *et al.*[159] This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated.

Implication for optimal primary care for OA GPs vary in the extent to which they deliver identifiable components of optimal care for OA, but training programmes can substantially improve their competence.

Study 2.3: development of the content of a model osteoarthritis consultation with a practice nurse and the training package for nurses to deliver it

Parts of this section have been reproduced from Healey *et al.*[160] © The Author(s). 2016 This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated.

Introduction/background

Evidence suggests that practice nurses are the health-care professionals who are most likely to provide self-management support for patients with chronic disease and, therefore, may be best placed to offer core treatments recommended in the NICE OA guidelines. The MOSAICS studies are the first to develop a system for delivering these core recommendations in the form of a linked GP–practice nurse model of care for OA patients.

Aim

To describe the development and content of the nurse-led component of optimal primary care for OA and the training package that aimed to provide practice nurses with the knowledge and skill set needed to enable self-management support for patients with OA in line with the NICE guidelines.

Methods

A rapid review of the literature and a mapping exercise of OA and chronic disease self-management programmes was conducted. Common elements were identified and, along with the NICE OA guidelines,[15] the content of the model consultation by the practice nurse (when referred patients with OA by the GP) was developed. Initial feedback on content was obtained from a practice nurse advisory group and the OA RUG.

Once the content was agreed, a training package was developed to provide the nurses with the knowledge and skill set needed to deliver the content. The training package was piloted with a mixed group of five nurses (a rheumatology nurse and a practice nurse, who were both also researchers, and three non-researching practice nurses) and results of the pilot used to refine the training package prior to delivery to all nurses in the intervention practices of the cluster RCT.

The training package was evaluated through (1) trainer reflections, (2) 'real-time' observations and suggestions from trainers and trainees and (3) formal daily evaluation by trainees.

Results

The content of the nurse-led OA service was finalised by the advisory group and the RUG as consisting of four 20- to 30-minute appointments over 3 months, following initial referral by the GP. The training package was agreed on as a 4-day programme.

The outcome of the development of intervention and training was that nurses were to focus on encouraging a patient-centred approach, the use of an OA guidebook, goal-setting, pain management and the core NICE recommendations (information and advice, exercise and physical activity, and weight management).

Evaluation of the pilot training resulted in theoretical material being reduced substantially, allowing more time during training to be dedicated to experiential elements (role play and simulated patient sessions). It was also suggested that an experienced facilitator should be involved throughout the training and that more time should be dedicated to practising the joint examination and to demonstration of exercise.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

The training programme was changed to reflect these findings of the pilot study.

Conclusion

Traditionally, OA management is not seen as a high priority for primary care and patients believe that little can be done. The NICE guidelines[15] highlight the therapeutic gains of positive self-management; practice nurse care and the training for this was developed and tailored to deliver and support these core messages.

Study 2.4: evaluation of practice nurse training to support osteoarthritis self-management

Parts of this section have been reproduced from Healey *et al.*[160] © The Author(s). 2016. This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated.

Introduction/background

Evaluation of the practice nurse training package was undertaken in a number of ways. All nurses who attended the training were asked to complete three evaluation questionnaires (pre training, 1-month post training and 6 months after initial intervention delivery).

In addition, the same pre-training questionnaire was given to practice nurses from the control practices to provide a baseline comparison. However, this group of nurses otherwise had no training as part of the trial and there was no special nurse-led referral service or OA clinics in the control practices.

The pre training and 1-month post-training questionnaires are presented here. Two other components of evaluation of the practice nurse training are considered in later sections of this chapter (see *Study 4.2: effect of the MOSAICS programme on patient-reported outcomes*, *Study 5.2: experiences of model consultations – patients*, and *Study 5.4: experiences of the model consultation – nurses*.) These were:

1. Patients recruited into the cluster RCT in both intervention and control practices were asked to complete a questionnaire about their experience and some of the questions were relevant to patients' experiences of the nurse-led clinics in the intervention practices.
2. Direct observations of the practice nurse-led clinics in the intervention practices during the cluster RCT were conducted and one-to-one interviews were conducted with a subsample of patients and practice nurses from the intervention practices.

Aims

To measure baseline knowledge and confidence of practice nurses in providing self-management support for patients with OA, and to examine the impact of an OA self-management training programme designed for practice nurses on the nurses' knowledge and confidence.

Methods

Practice nurses ($n = 25$) from all eight participating general practices in intervention and control arms of the MOSAICS trial were invited to complete a pre-training questionnaire to determine their knowledge and confidence regarding the management of OA using a 5-point Likert scale (higher scores = greater knowledge/confidence) and the practitioner self-confidence (PSC) scale (score range 4–20, > 7 = low confidence).[161]

Nurses ($n = 9$) from the practices that were randomised to deliver the trial intervention then attended a 4-day training programme designed to enable delivery of the core NICE recommendations (information and advice, exercise and physical activity, and weight management) using a patient-centred approach, the patient-designed OA guidebook and goal-setting to support self-management. Everyone who attended the training programme was asked to complete a post-training questionnaire.

Results

Pre-training questionnaire Of those invited to participate at baseline, 21 (84%) of the 25 nurses completed the initial questionnaire. All nurses were female, average years qualified was 25.8 years (SD 10.8 years), 9.5% were further qualified as nurse practitioners and some had experience or training in musculoskeletal medicine (23.8%) or orthopaedics/rheumatology (33.3%).

In terms of knowledge, only 4.8% scored ≥ 4 on the Likert scale when asked how much they had heard or read about the NICE OA guidelines. All nurses reported very low confidence on the PSC scale (16.1 SD 3.0) regarding the decisions needed when caring for patients with chronic joint problems, with 67% reporting no confidence in examining peripheral joints.

One-month post-training questionnaire Of the nine nurses who undertook the training, eight completed the post-training questionnaire. Training was associated with an increased knowledge of OA, with all nurses scoring ≥ 4 when asked how much they had heard or read about the NICE OA guidelines. Confidence also improved statistically significantly, with scores improving by 8.6 points on average ($p < 0.001$). Similarly, there was a clear shift post training in the proportion of nurses feeling OA management was part of their role (from 4.7% before training to 62.5% afterward scoring ≥ 4).

Conclusion

National Institute for Health and Care Excellence recommend that health-care professionals should support patients with OA to self-manage their condition. The results of this study in a small sample of nurses involved in the cluster RCT suggest an important gap between what is recommended and what nurses feel they can currently provide in terms of OA management. The development of a practice nurse training programme provided a systematic approach for developing the nurse contribution to delivery of optimal primary care for OA. Evaluation of the training suggests that it does contribute to practice nurses feeling more knowledgeable and confident in supporting patients to manage their OA more effectively.

Implications for primary care

Osteoarthritis appears to be an area in which practice nurses feel a lack of knowledge and confidence. Training for nurses in how best to support patients to self-manage their OA provides one route to meeting NICE guideline expectations of health professional engagement in OA care.

Study 2.5: development of the opportunistic osteoarthritis consultation with health-care professionals

Parts of this section have been reproduced with permission from Finney A, Porcheret M, Grime J, Jordan KP, Handy J, Healey E, *et al.*[144] Defining the content of an opportunistic osteoarthritis consultation with primary health care professionals: a Delphi consensus study. *Arthritis Care Res* 2013;**65**:962–968. © 2013 by the American College of Rheumatology.

Delphi consensus study

Introduction/background There is evidence that many older people with disabling joint pain do not mention this problem to health-care professionals or, if they do, it is in passing or as a lower priority than other conditions. Therefore, as part of developing an optimal approach to OA, there is a need to consider the whole primary care team, as any member of the team may encounter a patient with joint pain in the course of consultations about other conditions. This idea of such 'opportunistic consultations' about joint pain and OA is in addition to the model OA consultations with GP or practice nurse.

Aim The aim of this study was to define the content of a 'model' opportunistic primary care consultation for OA with any health-care professional in the primary care team.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Methods A two-round Delphi consensus exercise was conducted with four expert groups: GPs ($n = 30$), practice nurses ($n = 19$), allied health professionals ($n = 37$) (physiotherapists, occupational therapists, podiatrists, community pharmacists) and lay participants ($n = 18$).

Members of Arthritis Care, the NICE OA guideline development group and researchers at the Arthritis Research UK Primary Care Centre, Keele University, Keele, UK, helped to develop the potential consultation tasks. An ideas generation round and a two-round Delphi postal consensus study allowed participants to rank the importance of tasks for an opportunistic consultation.

Using the Calgary–Cambridge consultation framework,[162] the consultation tasks were developed into 35 statements. Participants in the consensus exercise were asked to rate the importance of the statements to include in the consultation. Agreement was defined as endorsement by ≥ 80% of participants across the groups.

Results There was a 50% response rate to the two-round postal exercise ($n = 52$). The response to the Delphi exercise varied between groups: GPs ($n = 11$, 37%), practice nurses ($n = 9$, 47%), allied health professionals ($n = 16$, 43%) and lay participants ($n = 16$, 89%).

The ideas generation round formulated 35 potential consultation tasks. Consensus was reached on 12 tasks for an opportunistic OA consultation using a ≥ 80% level of agreement across all groups.

Consensus was gained on three tasks using a 100% level of agreement across all groups: the health-care professional asks the patient (1) how things are going with their OA, (2) the type and amount of pain they have and (3) whether or not they are taking regular analgesia. In all, 12 consultation tasks were defined at a ≥ 80% level of agreement across all groups.

Conclusion In a Delphi study to define the content of an opportunistic primary care OA consultation, 12 consultation tasks provided the content of a comprehensive consultation. Three of these tasks with 100% agreement could be adopted in any multidisciplinary consultation for OA in primary care. These three tasks (a general enquiry about the condition, a question about the type and amount of pain the patient has and asking whether analgesia is being taken) form a core set of questions considered important by both lay and health professional groups in any opportunistic consultation with a patient who presents with OA or OA-related symptoms in primary care.

Implication for primary care Consensus between health-care professionals and patient groups suggest that a set of three simple enquiries could provide a practical guide for an opportunistic OA consultation:

1. General enquiry about the joint pain or OA.
2. How much pain is there and what it is like?
3. Are painkillers being used?

Dissemination workshops

Engagement and training of the wider primary care team in the intervention practices of the cluster RCT was achieved through a series of four workshops.

Aim To raise awareness of the MOSAICS model of primary OA care with allied health professionals in primary care linked with the intervention practices and to enhance consistency of care and continuity of care.

Methods Participants were all members of the multidisciplinary team linked to the four intervention practices. Representatives of allied primary health-care services linked to intervention practices were invited to four dissemination workshops facilitated by a nurse.

Results A total of 42 allied health professionals, one practice manager and two RUG members attended the four meetings. Community pharmacists were represented by nine participants, physiotherapy by 16 participants, occupational therapy by 10 participants and podiatrists five participants. There was one chiropodist and one health visitor.

The dissemination was undertaken by one of the Keele nurses responsible for running OA referral clinics in the intervention practices and supported by the research team. A video of a simulated consultation with the GP was shared. Information on the NICE guidelines,[15] the OA guidebook[163] and Arthritis Research UK 'Hands On' for OA were all distributed.[163,164] A mapping exercise of local services was performed and practical issues discussed.

Allied health-care professionals attending the workshops expressed their understanding of the potential value of the service and were willing to support patients referred to them as part of the new service or to refer patients back to the GP to enter the new service if they had come in by another route. Participants were keen to engage in feedback of the outcome of the programme and in future development of OA care.

Conclusion Good multidisciplinary representation in facilitated workshops produced a mapping of services for opportunistic consultations that might occur in sequence with the new model of care in the MOSAICS intervention practices.

Group 3 studies: development and evaluation of novel methods to measure quality of primary care for osteoarthritis patients

Outcomes in the cluster RCT were of two general types: measures of the implementation of good quality primary care for patients with joint pain and OA, and measures of the clinical effectiveness and cost-effectiveness of such care. Group 3 studies concerned the first type, namely measures of the quality of care delivered and received.

Measurement of the implementation of good-quality care used two approaches aligned with two of the populations of observed patients described in the introduction to this chapter (also see *Figure 7*):

1. Practice-level analysis of the total consulter population: all persons aged ≥ 45 years who were registered with one of the participating practices and who had consulted during the trial period about joint pain or OA. Outcomes used anonymised medical record data and included development of a novel computerised template to record items of care delivered [see *Study 3.1: the osteoarthritis quality-of-care template (the 'e-template')*].
2. Patient-level analysis of the total consenting population: the subgroup of the total consulting population, representing all persons responding to the baseline survey and who reported joint pain in the survey, who consented to future involvement and who subsequently consulted during the intervention period with joint pain or OA. Outcomes used individual self-reported data and included development with the OA RUG of a patient questionnaire for reporting experience of the consultation with the GP (see *Study 3.2: developing patient-reported outcomes with the osteoarthritis research user group*).

Study 3.1: the osteoarthritis quality-of-care template (the 'e-template')

Parts of this section have been reproduced from Edwards *et al.*[132] © The Author 2014. Published by Oxford University Press on behalf of the British Society for Rheumatology. This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com; and from Edwards *et al.*[145] This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 3.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/3.0/.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Introduction/background

Study 3.1 is concerned with the outcomes measured in the total consulting population, the whole registered populations of the participating practices aged ≥ 45 years who consulted about joint pain or OA during the trial period. Outcomes drawn from anonymised medical record data potentially represent indicators of the quality of care being provided by the practices. However, the number of such indicators routinely collected are limited and one aim of this study was to create and test additional markers that could be collected during the course of every OA consultation.

These markers (routine and new) were to be collected in all practices in the cluster RCT (intervention and control) in order to use them as an outcome measure for the trial, and were to be measured during a baseline period pre intervention as well as post intervention.

A novel method of collecting the new additional markers was developed in the form of a computerised pop-up 'e-template' for completion by the GP during a consultation with any patient aged ≥ 45 years who consulted about joint pain or OA. Such e-templates can be routinely triggered to appear on the computer screen and may act as aide-memoire, stimulus to action and method of recording that action.

However, although there is evidence that adherence to recommended processes of care may be improved through the use of such templates, there is no direct evidence for this effect of their use in OA care. Therefore, it was important to establish before the main trial began whether or not the use of such templates on their own, without the additional interventions described in the other group 2 studies above, might affect the quality of care, as the template was being used in all eight practices to collect outcome data for the trial. This was carried out by investigating the before-and-after effect of the e-template in all practices on a set of measures of the process and quality of care for patients with OA ('routine measures') prior to trial randomisation, also identified from anonymised routine primary care medical records but separate to, and not dependent on, the new e-template ('template measures').

Aims

1. To identify indicators of the quality of primary care for OA, which were also process measures of the implementation of care for patients with OA in general practice.
2. To develop and evaluate a pop-up computerised template (the 'e-template') as a method for improving recording of those indicators not already routinely captured in the medical record, specifically:

    i. to test the feasibility of its use in daily practice
    ii. to describe the quality of care as measured by the template
    iii. to determine the impact of incorporating the template into practice on other aspects of care already recorded routinely in primary care.

Methods

A systematic review to identify quality indicators for OA, applicable to primary care, was undertaken.[145] The outputs from the systematic review were used as the basis for developing methods for assessment and monitoring of the quality of care for OA in general practice.

The indicators identified in the systematic review were divided into:

- 'routine measures' – those which could be measured using routinely recorded information from consultations (e.g. prescriptions)
- 'template measures' – those which required a change to usual consultation recording practice (e.g. recording of the provision of advice). For this group of indicators, an e-template, designed to 'pop-up' during a consultation on the GP's computer screen, was developed, tested and implemented in all eight MOSAICS study practices prior to randomised allocation of the practices to intervention or control.

The recording template was set to be triggered automatically in all study practices by entry of one of a set of Read codes for OA or joint pain likely to reflect OA in an older patient population of all patients aged ≥ 45 years consulting and identified with one of the codes (see *Box 2*). The set of Read codes was identified by a panel of practising GPs and appears in www.keele.ac.uk/mrr,[140] Jordan *et al.*[108] and Edwards *et al.*[101] The template was installed and training was given regarding its use in all eight study practices.

An anonymised data set containing all routinely recorded data and all information entered into the templates was downloaded from the primary care computer systems to cover the period from 12 months before the template installation to 6 months post installation.

Informal feedback was requested from the practices about their views and experience of using the template at the end of the 6-month baseline period of template installation. In addition, formal feedback was investigated as part of the qualitative study integrated into the main trial through interviews conducted with GPs, nurses and other practice staff, and described in group five studies.

Analysis

1. Feasibility of using the template in primary care was measured by the frequency of its use and the extent of its completion during the 6-month post-installation period (baseline phase).
2. The pre-trial quality of care in the eight practices was measured by the pattern of template-derived indicators during the same baseline phase.
3. The impact that the installation of the template had on quality of care was estimated by comparing the pattern of non-template routinely recorded items in the primary care database during the 12 months prior to template installation with the 6 months after installation.

Results

A total of 25 different indicator themes identified in the systematic review were rated for their feasibility and applicability to UK primary care. This resulted in 15 indicators, informed by the NICE OA management guidelines,[15] being considered for inclusion in the study as process-of-care quality markers.

The template was designed to capture eight core treatment indicators which could not be measured from the routine record (pain assessment, functional limitation assessment, topical NSAID use, paracetamol use, OA information given, weight loss advice, exercise advice, and consideration of physiotherapy referral). A single-page e-template (10 items, eight quality indicators) for the Egton Medical Information System (EMIS) clinical information technology system was designed. This is summarised in *Appendix 2* (see *Items on the MOSAICS e-template*). The system was such that entry of one of the joint pain and OA codes immediately registered an administrative code that the template had been fired; a separate code was then entered for each item answered by the GP or other health professional on the e-template screen. If there were no coded entries for answers but there was a code indicating the template had been fired, this meant that the GP or consulting health professional had not used the template on that occasion.

The other measures of quality ('routine measures') could be identified through routine medical record data separate to the e-template and its use (prescribing of paracetamol, topical NSAIDs, opiates, oral NSAIDs with or without gastroprotection using a proton pump inhibitor, relevant radiography use, and onward referral to selected specialities including exercise referral or physiotherapy, occupational therapy, weight loss programmes, orthopaedics, pain medicine, and rheumatology).

The main findings from analysis of the baseline assessment of quality of care using the new e-template were that:

- The template is feasible, triggering on 93% of 1851 eligible patients during the 6 months baseline phase. Out of 1730 patients triggering the template, 66% had at least one template item achieved and 20% achieved all template-derived indicators.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

- Quality of care varied by indicator, with good achievement for pain (63%) and function (62%) assessment; consideration of physiotherapy was the least well-achieved at 36%.
- There was considerable variation in management behaviour between individual clinicians.
- Patients who had a formal OA diagnosis entered into the Read code system, rather than a symptom-based joint pain code alone, achieved more quality indicators.
- Overall recorded quality of care for those elements common to previous studies on this topic was similar to that reported in those earlier studies.[165]

The analysis of the separate routinely recorded measures of primary care OA management showed that:

- Morbidity coding of OA was stable before and after template installation.
- There was a statistically significant increase in prescribing of recommended first-line analgesics, paracetamol and topical NSAIDs in the 6 months after template installation compared with the 12 months before the template was inserted.
- There was a statistically significant increase in weight recording after template installation, which then fell but remained above pre-template levels.
- There were no changes in other management actions which appear in routine medical records separately to the e-template recording.

Informal feedback from the practices covered a range of opinions. In general, the GPs involved reported that the template was useful initially, although some felt that they then became 'attuned' to it and no longer required the template to trigger in order to record the necessary information. Others reported it to be useful but wanted some modification to reduce the frequency of triggering, and some identified a need for a wider spread of quality markers to be collected. Some reported feeling wearied by the number of such standard pop-up screens and expressed reluctance to continue to use the template. Further feedback will be reported in the qualitative study (see *Study 5.1: implementing a complex intervention in practice – an evaluation using a theoretical framework*, and *Study 5.3: experiences of model consultations: general practitioners*).

Conclusion

A template such as the one developed and evaluated here does appear to provide a useful basis for monitoring, measuring and auditing core primary care for OA patients and for measuring implementation of care in general practice. In addition, the template appeared to prompt improvements in recommended initial management of OA, as measured by other routinely collected information.

The introduction of the e-template did not appear to affect or change the way in which morbidity codes were selected and allocated to older patients presenting with joint pain.

Feedback from doctors and nurses at the participating practices indicated a general level of satisfaction with the practicality and usefulness of the e-template, but individual practitioners expressed concern at the workload and desirability of its long-term use. The longer-term effect of the templates was therefore planned through future analysis of the control practices in the cluster trial, which continued to use the template alone for a further 12 months.

Implications for primary care

A standard template on a GP's computer screen, routinely triggered when a consulting code is entered for joint pain or OA, offers a feasible way to collect information on quality of care not routinely available from the medical records.

Adoption of such a template appears to improve quality of OA care, but ways to improve its integration into the consultation and to avoid 'template' overload are important issues to be addressed by future research.

Study 3.2: developing patient-reported outcomes with the osteoarthritis research user group

Parts of this section have been reproduced from Blackburn *et al.*[166] © 2016 Blackburn *et al.* Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated.

Introduction/background

There was a need to study process measures in the cluster trial regarding whether or not better quality of care for OA patients in primary care had been achieved by the new intervention. The previous section has described the introduction of the OA consultation template, which had the potential to both promote and record GP activity. However, such templates cannot measure the patient experience of care and so a set of self-report patient questions was developed to identify the patient experience and perception of the OA care that they had received. The OA RUG group led the work on this with members of the research team.

Aim

To identify and select items for the MOSAICS self-reported quality-of-care questionnaire.

Methods

The pool of potential indicators was provided by the systematic review by Edwards *et al.*[145]

Patients with OA who were members of the OA RUG group were invited to a consultation meeting with two members of the research team (JE and KD). The RUG members were asked to identify indicators of quality of care perceived as relevant to primary care consultations. Each RUG member was asked to prioritise their five most important indicators. RUG members also helped to develop the wording and scoring of the patient-completed questionnaire that would draw on their selected items (the MOSAICS OA quality questionnaire).

A draft of the questionnaire was prepared following the meeting and reviewed and amended by RUG members.

For validation purposes, the draft instrument was compared with a new researcher-generated questionnaire developed and validated for secondary care by Østerås *et al.*[146] for a Norwegian population [Musculoskeletal pain in Ullensaker Study Osteoarthritis Cohort Quality Indicator (MUST OA-QI) questionnaire]. The RUG fed back their views on the MUST OA-QI questionnaire and how it compared with the MOSAICS questionnaire, both individually and in a follow-up group meeting.

A teleconference was then held between representatives of the RUG group and researchers currently using the MUST OA-QI questionnaire to discuss the similarities and differences. A final RUG meeting summarised the process.

Results

Six RUG members with OA (three men and three women) attended up to five meetings. From 30 quality indicators initially identified, 20 were prioritised as important by RUG members. These covered pain self-management, medication usage and side effects, help with daily activities, information about OA, help with improving QoL, and level of support.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

From this initial list, the RUG members helped to develop and refine the wording of 14 items and determine the scoring system for the MOSAICS quality questionnaire. The scoring system was the same in the MUST OA-QI questionnaire of 17 items. RUG members determined that both questionnaires contained the same or similar concepts. They felt that all of the Musculoskeletal pain in Ullensaker Study Osteoarthritis (MUST) indicators were important to OA patients, but some of the terminology from back translation was not applicable for UK primary care. The MUST questionnaire included two items relating to weight management that were excluded from the MOSAICS questionnaire because they were being separately recorded on the OA e-template.

Conclusion

Informed RUG members and researchers can agree on important indicators of the quality of care for OA. Furthermore, while RUG involvement in MOSAICS established the relevance of items and refined wording, the Norwegian MUST researchers had previously tested the validity and reliability of the instrument ready for use. The coincidental development of the two questionnaires provided the opportunity to compare the generalisability of the items from two European countries and to explore the comparative construct validity of the MOSAICS draft tool. The Keele OA RUG group and researchers produced similar self-reported measures of quality of OA care to the group of researchers from Norwegian secondary care. Patient involvement ensured the relevance of quality indicators to patient experience and the refining of questionnaire wording.

Group 4 studies: clinical effectiveness and cost-effectiveness of the MOSAICS programme to optimise primary care of osteoarthritis

Once the baseline population survey and the development studies had been completed, including the evaluation of introducing an e-template in all eight participating practices described in *Study 3.1: the osteoarthritis quality-of-care template (the 'e-template')*, randomisation of the eight practices and introduction of the MOSAICS model of care into the four intervention practices took place.

There were two main quantitative studies contained within the main trial, reflecting the twin aims to investigate:

1. implementation of the MOSAICS model of care using practice-level analysis of anonymised consultation data (study 4.1) and
2. clinical effectiveness and cost-effectiveness of the MOSAICS model using patient-level analysis of self-report follow-up data and medical record data (studies 4.2 and 4.3).

Qualitative studies addressed implementation questions through interview studies with both patients and health-care professionals (group 5 studies).

Study 4.1: implementation of the MOSAICS intervention – practice-level effect on quality of care for patients with osteoarthritis

Parts of this section have been reproduced from Jordan *et al.*[167] © 2017 The Authors. Published by Elsevier Ltd on behalf of Osteoarthritis Research Society International. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Introduction/background

The introduction of the e-template, prior to randomisation, for recording items of care delivered during the OA consultation in all eight participating study practices, established a baseline level of achievement of quality-of-care indicators for OA and demonstrated, through a before-and-after analysis, an improvement in some non-template routinely recorded items of management of OA [see *Study 3.1: the osteoarthritis quality-of-care template (the 'e-template')*].

The next questions to be addressed were those from the post-randomisation period and were concerned with the implementation of the MOSAICS model of care for all patients consulting with joint pain or OA in

the four intervention practices. The main question to be addressed was whether or not introduction of the MOSAICS model of care resulted in further change in quality-of-care indicators (as evidence of putting NICE guidance into practice) for patients consulting with joint pain or OA.

Aims

To determine whether or not implementation of the MOSAICS model of care post randomisation (patient information booklet, model GP consultation, practice nurse-led series of OA follow-up consultations, training of all practice staff in opportunistic OA consultations) was successful at practice-level as measured by:

1. process measures of the implementation of the MOSAICS model of care in the four intervention practices
2. improvements in the recorded management of all patients aged ≥ 45 years consulting with joint pain and OA (e-template and routine) in the four implementation practices compared with the four control practices.

Design

Cluster RCT in eight general practices (four intervention and four control) with practice-level analysis.

Methods

Training period After the 6-month run-in period, during which all eight practices used the e-template when any adult aged ≥ 45 years consulted for joint pain and OA, the eight practices were randomised to intervention or control. All practices continued to use the template but intervention practices were then trained during a 3-month training period to deliver optimal primary care for OA according to NICE guidance.

Optimal care The MOSAICS model of care for patients with OA, developed as described in the group 2 studies above, was introduced into the four intervention practices, with GPs, nurses and practice staff in these four practices receiving the training over a 3-month period. Copies of the OA patient guidebook were supplied to all intervention practices to reinforce verbal information given by GPs and practice nurses and to provide additional information about OA and its management and about local resources to support self-management. The GPs were trained in a series of four in-practice sessions (3 × 2 hours; 1 × 1 hour) to deliver the model OA consultation, including provision of the OA guidebook and the offer of a series of up to four follow-up consultations with the practice nurse. The practice nurses were trained across 4 days by a programme of education activities, workshops and simulated consultations in the delivery of OA care to those patients who were referred to the follow-up clinics by the GP and who chose to attend.

In contrast to the comprehensive training programme developed for the GPs and practice nurses, one-off workshop events were used to disseminate the core set of consultation questions for use in opportunistic consultations about OA by any member of the extended primary care team linked to the four intervention practices.

Usual care No further training or resources were provided for the control practices in caring for patients consulting with joint pain and OA, apart from continuing support for the use of the e-template as the method for monitoring and auditing quality of care.

Follow-up period All routinely recorded medical record data for patients aged ≥ 45 years were downloaded and anonymised from all eight participating general practices for a minimum of 12 months after the training period ended (i.e. after the end of the 3-month training, a time point designated as the 'start of the trial'), as were data on quality indicators for the management of patients with OA recorded in the templates. This included 4 million consultation records and 6 million prescription records, as well as recorded information on registration status, investigations, tests and referrals.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Practice-level outcomes

1. Process measures of implementation of the MOSAICS model of care in the intervention practices included the number of patients who saw the practice nurse in one or more OA follow-up clinics.
2. Effect of implementation of the MOSAICS model of care on the delivery of NICE guidance in the intervention practices compared with control included any differences after the start of the trial in:

    i. the total number of patients coded as consulting with joint pain and OA and the separate numbers coded with joint pain or OA, expressed as a proportion of the total registered population of ≥ 45 years in the practices
    ii. the proportion of patients coded as consulting with joint pain or OA who had at least one item on the e-template completed, and the proportion with all items completed
    iii. the proportion of patients with joint pain or OA and with at least one item completed on the e-template who met each separate criterion for quality of care on the e-template
    iv. the proportion of patients with joint pain or OA who received routinely recorded items of care not dependent on e-template completion.

Analysis

Items of recorded management and achievement of quality indicators were calculated for all patients at the eight practices aged ≥ 45 years who were recorded as consulting with OA or joint pain during the first 6 months of follow-up after start of the trial. This was done as follows.

'E-template measures' Quality indicator information recorded through the template (fired every time an OA or joint pain code was entered) was extracted for a 120-day period following an individual patient's first consultation after the start of the trial. This information included assessment of pain and functional impairment, provision of information, advice about exercise and weight loss, and advice and consideration of pharmacological management involving paracetamol and topical NSAIDs. Achievement of an indicator was defined as an item recorded in the template as having been performed, considered, offered or deemed not appropriate. Further analyses also considered specific elements of the template, namely the recorded provision of written information on OA, written exercise advice, whether or not paracetamol was offered and whether or not topical NSAIDs were offered.

Additional practice nurse resource Uptake of the additional practice nurse resource in intervention practices was analysed to explore the extent to which this resource may have contributed specifically to quality of care in the intervention practices. The proportion of all patients consulting in the intervention practices who attended the practice nurse OA clinics was calculated and e-template completion was compared between those who did and did not see the nurse.

'Routine measures' Routinely recorded management was identified within 14 days of any OA consultation (including the first) taking place within 120 days of the first such consultation. This management covered prescribing (paracetamol, topical NSAIDs, opioids, oral NSAIDs, weight loss agents), investigations (use of relevant radiography), and referral to selected specialities including exercise referral or physiotherapy, occupational therapy, weight loss programmes, orthopaedics, pain medicine, and rheumatology. Prescriptions for a proton-pump inhibitor in those prescribed an oral NSAID were also identified.

Differences between intervention and control practices These differences in achievement of quality indicators on the e-template and in routinely recorded management were adjusted for age, sex, whether the initial consultation was recorded as diagnosed OA or joint pain, and baseline level of management or achievement of the practices (i.e. during the baseline 6 months when the template was introduced prior to the start of the trial). Clustering of patients within clinicians was accounted for through the use of multilevel modelling, estimated using iterative generalised least squares with second order penalised quasi-likelihood approximation. Results are presented as ORs with 95% CIs.

Sensitivity analyses Sensitivity analyses restricted analyses to:

1. new consulters for OA or joint pain after the start of the trial (first such consultation since the introduction of the template and no OA or joint pain consultation for the previous 365 days)
2. those with a recorded diagnosis of OA (as opposed to a joint pain record)

and, for the analysis of quality indicators,

1. those with at least one recorded entry on the template under the assumption that if no entry was recorded, the GP was less likely to believe the patient had OA.

Adverse events

Recorded occurrence of 15 pre-specified adverse events for a median of 14 months after initial OA or joint pain consultation was also compared between arms.

Results

Ten general practices in Staffordshire and Cheshire were approached to take part and eight were recruited. *Table 13* shows practice characteristics according to status as intervention or control. Intervention practices had a mean total registered population size of 10,240 (SD 9174.8) compared with 6983 (SD 2060.7) for control practices. The intervention practices were more deprived on average than control practices. Despite these differences, the mean number of GPs per practice was similar: 6.0 (SD 6.1) for intervention and 5.5 (SD 2.9) for control practices, and the mean age of the GPs was similar.

Figure 7 shows the flow chart for the practice-level analysis of the trial. There were 33,726 adults in total aged ≥ 45 years who were registered with one of eight participating practices at the time of the baseline population survey: 18,835 in the intervention practices 14,891 in the control practices. During the trial period of 6 months, which began after the completion of training in the intervention practices, 1960 persons from this population base consulted at least once with joint pain or OA when one of the Read codes was entered by the GP on to the computer record: 1118 in the intervention practices and 842 in controls. This equates to a 6-month consultation prevalence for joint pain and OA of 5.78 persons aged ≥ 45 years per 100 in the participating practices (5.94 per 100 in the intervention practices and 5.65 per 100 in the control group). In the intervention arm, the mean age was 66.2 years (SD 12.3 years) and 59% were female, compared with control arm mean age 66.5 years (SD 11.9 years) and 61% female.

There was no change in the proportion of all consulters given an OA code rather than a joint pain code in the intervention arm in the 6 months before and after randomisation (43% before, 45% after). However, throughout both periods, these figures were different from those in the control practices (23% received an OA code in the 6 months before randomisation in the control practices; 29% during the 6 months after randomisation).

TABLE 13 Characteristics of the eight general practices participating in the cluster trial

Characteristic	Control ($n = 4$)	Intervention ($n = 4$)
Number of registered patients per practice, mean (SD)	6983 (2060.7)	10,240 (9174.8)
Practice Index of Deprivation, median (IQR)[a]	14,633.5 (4571.5–28,822.0)	9165.0 (2195.7–19,478.5)
Number of GPs per practice, mean (SD)	5.5 (2.9)	6.0 (6.1)
Age of GP (years), mean (SD)	42.8 (23.5)	42.2 (23.7)

IQR, interquartile range.
a The lower the index, the more deprived the area.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

e-template outcomes Successful firing of the template was similarly high in intervention and control practices – 1061 (95%) of those with a relevant code in the intervention group, 757 (90%) in the control group. However, in a substantial proportion of patients in both intervention and control groups, no items in the template were completed at all (40%).

As detailed in study 3.1 [see *Study 3.1: the osteoarthritis quality-of-care template (the 'e-template')*], there continued to be wide variation in baseline achievement of e-template quality indicators between individual clinicians and between practices, and this was reflected in baseline differences between the two trial arms. Although there were generally higher levels of achievement of indicators during the trial period in the intervention arm, the changes from baseline were not statistically significantly different from baseline levels when compared with the changes in the control arm (*Table 14*). There was also evidence of a fall in the levels of achievement between baseline and trial period for both intervention and control practices, although this was not apparent in those with at least one entry on the template (data not shown).

TABLE 14 Comparison between intervention and control practices of all consulters firing template, by (1) e-template quality indicators and (2) routinely recorded management

	Baseline period		Trial period		
Quality indicator	Intervention, *n* (%)	Control, *n* (%)	Intervention, *n* (%)	Control, *n* (%)	OR[a] (95% CI)
e-template quality indicators					
Number of consulters	981	749	1061	757	
Pain assessment	707 (72)	390 (52)	617 (58)	318 (42)	1.35 (0.58 to 3.14)
Function assessment	691 (70)	384 (51)	611 (58)	309 (41)	1.15 (0.49 to 2.71)
Topical NSAIDs	540 (55)	295 (39)	501 (47)	275 (36)	0.97 (0.48 to 1.95)
Paracetamol	625 (64)	349 (47)	554 (52)	284 (38)	1.42 (0.71 to 2.95)
Information given	578 (59)	274 (37)	554 (52)	268 (35)	1.34 (0.61 to 2.96)
Exercise advice	582 (59)	285 (38)	526 (50)	246 (32)	1.53 (0.75 to 3.13)
Physiotherapy referral	426 (43)	192 (26)	348 (33)	173 (23)	1.45 (0.61 to 3.40)
Weight loss advice[b]	325 (53)	159 (34)	341 (49)	136 (31)	1.24 (0.61 to 2.52)
1-plus indicator achieved	727 (74)	419 (56)	635 (60)	330 (44)	1.49 (0.65 to 3.43)
All indicators achieved	246 (25)	106 (14)	240 (23)	116 (15)	1.57 (0.70 to 3.50)
Routinely recorded management					
Number of consulters	1015	836	1118	842	
Paracetamol	164 (16)	155 (19)	241 (22)	117 (14)	1.74 (1.27 to 2.38)
Topical NSAIDs	267 (26)	194 (23)	327 (29)	186 (22)	1.21 (0.83 to 1.76)
Opioids	344 (33)	244 (29)	367 (33)	232 (28)	0.93 (0.75 to 1.17)
Oral NSAIDs	181 (18)	119 (14)	176 (16)	137 (16)	0.78 (0.52 to 1.16)
Any analgesic	645 (64)	484 (58)	711 (64)	460 (55)	1.20 (0.92 to 1.58)
Proton pump inhibitor	63 (35)	27 (23)	69 (39)	50 (36)	0.92 (0.43 to 1.98)
Weight record	278 (27)	154 (18)	309 (28)	144 (17)	1.36 (0.80 to 2.33)
Referral	233 (23)	139 (17)	252 (23)	175 (21)	1.00 (0.72 to 1.39)
Physiotherapy referral	90 (9)	35 (4)	111 (10)	19 (2)	5.30 (2.11 to 13.34)
Radiography	250 (25)	22(3)	163 (15)	47 (6)	0.45 (0.12 to 1.72)

a Adjusted for age, sex, coded OA or joint pain, practice level of achievement in baseline period and clustering by clinician (reference is control group).
b In those recorded as overweight.

However, there were observed changes in the core NICE recommendations which were substantially and statistically significantly different in the intervention group and the control group. In the intervention group, supplying written information increased from 4% of patients at baseline to 28% post intervention, exercise advice from 4% to 22%, and weight loss advice in overweight patients from 1% to 15%. These increases were more pronounced in those with an OA diagnosis (data not shown). In contrast, the difference in patient proportions receiving any of the three recommendations between baseline and follow-up was no more than 1%.

Comparisons of routinely recorded management are shown in *Table 14*. Prescribing of paracetamol increased from the baseline period in the intervention arm and decreased in the control arm. Adjusting for baseline prescribing, the intervention arm patients were statistically significantly more likely to receive paracetamol and to be referred to physiotherapy, and (non-significantly) less likely to be sent for a radiograph. There were no differences in the other routinely recorded management actions. Those with an OA diagnosis were additionally more likely to receive any analgesic and to have weight recorded in the intervention practices.

Practice nurse attendance Results of e-template quality indicator analysis in the four intervention practices by practice nurse-led clinic attendance are shown in *Table 15*.

Table 15 shows the results for the 1060 patients who consulted in the four intervention practices in the main trial period and were given one of the joint pain or OA Read codes. Most of these patients (79.2%)

TABLE 15 Details of e-template use in 1060 patients recorded as consulting with joint pain or OA in the intervention practices, by whether or not practice nurse-led clinic attendance occurred

Quality indicator	Coded as consulting but did not attend nurse-led clinics, $N = 840$ (79.2%)		Coded as consulting and did attend nurse-led clinics, $N = 220$ (20.8%), $n = (\%)$
	Template fired but no entries, $n = (\%)$	At least one template entry, $n = (\%)$	
Number of consulters	424	416	220
OA diagnosis	135 (31.8)	198 (47.6)	149 (67.8)
Core recommendations			
Written information	–	100 (24)	195 (89)
Written exercise advice	–	55 (13)	177 (80)
Written weight loss advice	–	29 (10)	75 (44)
Quality indicator achievement			
Pain assessment	–	398 (96)	218 (99)
Function assessment	–	392 (94)	218 (99)
Topical NSAIDs	–	316 (76)	184 (84)
Paracetamol	–	352 (85)	201 (91)
Information given	–	338 (81)	215 (98)
Exercise advice	–	309 (74)	216 (98)
Physiotherapy referral	–	215 (52)	132 (60)
Weight loss advice	–	193 (69% of those overweight)	147 (87)
At least one indicator achieved	–	416 (100)	218 (99)
All indicators achieved	–	132 (32)	107 (49)

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

did not attend the nurse-led clinic subsequently but for about half, the GPs bypassed the template (i.e. the template fired but the GP chose not to complete it), while for the other half, the GPs completed the templates although the patients did not attend the nurse-led clinic.

Among patients who had at least one item completed on the template, the pattern of items completed was compared between those patients who did and did not see the nurse. The main contrasts are (1) the higher proportion of patients who were provided with written (rather than verbal only) information, exercise advice and weight loss advice and (2) the higher proportion who met the criteria for discussion of physiotherapy referral (52%) in the group who saw the nurse.

Adverse events

There were 13% of the intervention arm and 11% of the control arm with a recorded adverse event. Just under half of recorded adverse events overall were falls. Differences between arms were small and the sole statistically significant difference between arms was for heart failure (1.5% intervention arm vs. 0.5% control arm). Heart failure was not associated with paracetamol or NSAID prescribing.

Discussion

Patients with no e-template entries Across both intervention and control practices, about 40% of patients who attended with joint pain and OA and received a code which fired the e-template had no template items completed. There are a number of possible explanations.

1. The pool of potential patients for whom the template would be fired was broad, given the wide range of eligible joint pain and OA Read codes. It is likely that, for some patients, the GP did not consider template completion relevant to the clinical presentation, as suggested in *Table 15* by the lower percentage of persons with a diagnostic OA code among those with no template entries. However, this still means that almost one-third of this group in the intervention practices were diagnosed with OA but the GP chose not to complete the template.
2. GP variability. There were known to be periods when locum cover resulted in GPs untrained in MOSAICS consulting with patients and the unfamiliar template may have been bypassed; immediate workload and other more pressing patient concerns might also have led to this outcome.

Neither explanation should have affected the main trial result, despite substantial differences between intervention and control practices in the proportions receiving a diagnostic code of OA in particular. However, the finding is relevant to dissemination: up to two out of five patients may have an uncompleted template due to patient inappropriateness or GP concerns.

e-template criteria for quality of care As occurred in the pre-randomisation phase [study 3.1, see *Study 3.1: the osteoarthritis quality-of-care template (the 'e-template')*], criteria for quality of care as measured by completion of the template were met for most indicators in a substantial majority of patients in both arms of the trial (pain assessment, function assessment, consideration given to topical NSAIDs and paracetamol, information given, exercise advice and weight loss advice for those who were overweight). The main difference between intervention and control practices post-randomisation was the substantially higher proportion of patients receiving written information in the intervention practices, although the point estimates for a number of criteria suggested higher but statistically non-significant levels of achievement in the intervention practices.

Within the intervention practices, GPs, once they decided to complete the template, appeared to do so for most items, regardless of whether or not the patients were going to see the nurse. However, the proportion recording that they are providing written (rather than verbal only) information, exercise advice and weight loss advice was higher in the group seeing the nurse, which suggests that these criteria are less achievable in a GP consultation, despite the availability of the patient information booklet in all intervention practices. Variability in the number of trained GPs seeing patients and in workload and priorities may also play a role.

A limitation of the study is that, for many e-template items, baseline variation between practices and clinicians was large, and intervention and control practices were already different in aspects linked to higher quality of care in the intervention practices, notably higher baseline rates of OA diagnosis among people aged ≥ 45 years presenting with joint pain. However, we have adjusted our results for these differences.

Routinely recorded management Despite the problems of the contrasting baseline proportions of patients with an OA diagnosis between intervention and controls, and the already high levels of template criteria achieved pre-randomisation in both intervention and control practices, there were statistically significant differences between intervention and control practices in objective measures of GP behaviour drawn from the routine medical records, namely a change in analgesic prescribing and an increase in physiotherapy referral (although prevalence of the latter was low). This does suggest that the MOSAICS model of care is shifting GP behaviour in the direction of better OA care.

Nurse clinic attendance as measure of MOSAICS implementation The full implementation of the MOSAICS model of care implied that patients would receive the booklet and see the practice nurse. Only one out of five patients consulting with joint pain or OA in the intervention practices attended practice nurse-led clinics. However, a substantially higher proportion of these patients received written information about NICE core guidance than patients not seeing the nurse. This suggests two things:

1. The main trial difference between intervention and controls (increase in provision of written information) was achieved substantially in the subgroup who saw the nurse despite this subgroup being a minority of the intervention group as a whole.
2. The MOSAICS intervention was not fully implemented.

The implication is that the nurse-led clinic offered further opportunities for template items and quality criteria to be completed at the time of nurse attendance, and that seeing the nurse was linked to full implementation of the MOSAICS method of delivering NICE core guidance including the booklet and opportunities to discuss exercise and physiotherapy referral. It is possible that this was selection on the part of the GPs; patients may be more likely to be referred to the nurse if the GP had already completed more e-template activity, but it is also likely to be related to additional activity and recording by the nurse.

Strengths and limitations One strength was the pre-randomisation introduction of the e-template in all practices. This meant that any initial effect of the e-template on the recording and performance of quality indicators of OA care should have been the same across all practices and independent of future intervention or control status. This enabled separation of the effects of the model OA consultation from those of the template itself. It also meant that the template could be used as a standardised format for collection of outcome data in both intervention and control practices. However, this came at the cost of not being able to estimate the effect of the model OA consultation in the absence of the template. Ideally, a third arm receiving neither template nor model OA consultation would have provided maximum information from the trial. However, this would have severely restricted the number of process-of-care measures available as trial outcomes and greatly inflated sample size requirements.

Although GPs in the control practices were of necessity engaged in the use of the template and their recording of care was influenced by this, they were not involved in any way in the development of the MOSAICS model OA consultation and its components. This development took place in general practices entirely outside the eight practices involved in the trial itself and its application in the trial was restricted entirely to the intervention practices.

One weakness is that we did not test or investigate the additional time taken by GPs to use the template, deliver the model consultation, introduce the booklet and arrange follow-up. We recognise that there will be challenges in delivering the intervention in time-constrained general practice and this is one issue that needs to be studied in future implementation.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Conclusion

This study provides evidence that the MOSAICS model of care was implemented to an extent that substantially increased the frequency of reported provision of written information to patients consulting with joint pain and OA in primary care. Most of this increase appeared to take place in association with at least one visit by the patient to the practice nurse OA clinic in one of the intervention practices.

The conclusion is that the MOSAICS intervention can be implemented and can be carried out to an extent that improves quality of care beyond that achieved by use of the template alone. However, not all items of care will be changed by the MOSAICS model of care and only a minority of patients saw the nurse.

The written patient-focused information resource available in the MOSAICS model of care is likely to have provided the basis for the improvement in written advice and information, but appears to have depended on a patient attending the nurse-led OA clinics.

The intervention does appear to alter clinician behaviour in practice, as measured by routine records, beyond changes achieved by the template alone (notably, increased paracetamol analgesia, increased physiotherapy referrals, and decreased radiography referrals).

Levels of achievement of quality indicators as measured through the template generally fell from baseline levels. Further work is needed to assess if familiarity with the template meant it was no longer needed in the longer term as a prompt for management behaviour or if this represents a decline in initial improved quality of care.

Study 4.2: effect of the MOSAICS programme on patient-reported outcomes

Parts of this section have been reproduced from Dziedzic *et al.*[168] © 2018 The Authors. Published by Elsevier Ltd on behalf of Osteoarthritis Research Society International. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Introduction/background

Aim

The primary aims of this study were to determine:

- patient-reported measures of the effectiveness of implementation of the MOSAICS model of care
- the clinical effectiveness of the model OA consultation in participants aged ≥ 45 years with joint pain and who were registered in general practice.

Design

Prospective recruitment of eligible individuals who had previously provided consent, conducted at the time of a consultation about joint pain and OA, into a follow-up using self-report questionnaires, within a cluster RCT in eight general practices (four intervention and four control).

Methods

Figure 7 illustrates the approach. A population survey was mailed to all adults aged ≥ 45 years who were registered with one of eight participating practices to establish the potential pool of individual trial participants in advance of randomisation of the practices to intervention and control. This was done through written consent for medical record review and future follow-up contact, which was contained within the questionnaire. All responders to the survey who had reported pain in at least one of four target OA sites (hand, hip, knee, foot) and who gave this consent had a tag inserted in their general practice medical record as an identifier. The tagged records identified eligible patients if and when they consulted with joint pain or OA.

After the 6-month run-in period, during which all eight practices used the e-template when any adult aged ≥ 45 years consulted for joint pain and OA, the practices were randomised to intervention or control. All practices continued to use the e-template but intervention practices were trained for a 3-month period to

deliver the MOSAICS model of care for implementing NICE guidance for patients with OA, while the control practices continued with usual care (described in *Group 2 studies: developing interventions to support self-management of osteoarthritis in primary care and to monitor the quality of care delivered*).

For the individual patient study, all patients aged ≥ 45 years who consulted with joint pain and OA (hip, knee, hand and foot) in both intervention and control practices during the 6 months of trial follow-up (i.e. from the 'start of the trial' following the 3-month training period for the intervention practices) and who had a consent tag inserted in their records, were recruited into the study. A follow-up questionnaire was mailed to all recruited persons immediately following the index consultation (0 months) and at 3, 6 and 12 months following index consultation.

There were two types of end point measured in the individual patient study.

Process measures of implementation of the MOSAICS model-of-care

1. Uptake of NICE recommendations was measured using patient self-reported questionnaires:

    i. A previously validated questionnaire for OA patients enquiring about their use of services and self-management [the modified Knee Pain Screening Tool (KNEST)[169]], with a selection of items that measure care related to NICE guidance topics.
    ii. The osteoarthritis quality indicators (OA-QIs) questionnaire for UK primary care, the development of which is described in study 3.2 (see *Study 3.2: developing patient-reported outcomes with the osteoarthritis research user group*).

2. Additional process measures of the implementation of the MOSAICS model were provided by:

    i. a measure of fidelity of the intervention – the proportion of patients in the intervention practices visiting the nurse after their index consultation with the GP in the individual-level analysis [similar to its use in the whole population analysis described in study 4.1 (see *Study 4.1: implementation of the MOSAICS intervention – practice-level effect on quality of care for patients with osteoarthritis*)
    ii. a set of questions about what had happened in the consultation, developed as part of the GP training programme (see *Group 2 studies: developing interventions to support self-management of osteoarthritis in primary care and to monitor the quality of care delivered*)
    iii. Physical Activity Scale for the Elderly (PASE)[170] and International Physical Activity Questionnaire (IPAQ)
    iv. enablement as measure of confidence to self-manage.

Measures of clinical effectiveness of implementing the MOSAICS model of care measured at 6 months after the index consultation The primary outcome was general physical health (SF-12 PCS[171]). Secondary outcomes were joint pain intensity and pain self-efficacy.

Analysis

1. Baseline comparability
   The baseline survey responder and practice characteristics were compared between practices according to their intervention control status. The survey characteristics of the subgroup subsequently recruited into the individual trial follow-up were then also compared between intervention and controls.
2. Process outcomes
   The self-report items were compared between intervention and control groups using logistic regression. The proportion of patients seeing the nurse was analysed within the intervention group according to information in the 3 month follow-up questionnaire as it was judged that this would give adequate time for all patients who were referred to the nurse to have had the opportunity of at least one visit. Among those who did visit the nurse, a case report form (CRF) was completed for each patient as a measure of the content of care provided by the GP and the practice nurse.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

3. Clinical outcomes
   A linear mixed model was used to compare intervention and control group participants with respect to the primary clinical outcome (SF-12 PCS). A three-level mixed model was fitted to test for the effect of the intervention from baseline across all follow-up points, taking into account clustering by practice, participant and repeated measures follow-up.

A per-protocol complier average causal effect (CACE) sensitivity analysis was also performed. The purpose of this was to provide an unbiased estimate of the treatment effect in the subgroup of patients who received treatment administered per protocol in the intervention arm. This was defined as all those who saw the practice nurse as a follow-up to the GP consultation.

Sample size calculation This was based on the primary clinical outcome (SF-12 PCS at 6 months post consultation). Published estimates of the minimal clinically important difference for the SF-36 (version 2) puts it at 2–4 points on the PCS subscale, which has a normalised population SD of 10 points. Hence, a difference between groups of 0.3 effect size was calculated as the threshold for demonstrating difference between control and intervention groups in the MOSAICS studies.

A total of 500 participants had to be recruited at consultation to allow for a drop-out rate of 20% and to detect an effect size of 0.3 with 90% power at 6 months, given a 5% two-tailed significance level. Randomisation was by practice, so the sample size calculation was inflated to correct for an intracluster correlation coefficient (ICC) (adjusted ICC of 0.005). Varying practice size recruitment was taken into account (including coefficient of variation of 0.5)[131] and included (× 0.67 and × 1.25, respectively for intervention and control practices) adjustments for repeated-measures design.

Results

The flow chart for recruitment to the study is shown in *Figure 7*. There were 15,083 persons who completed and returned their questionnaires in the baseline survey out of a total of 28,443 adults aged ≥ 45 years registered with the practices and considered eligible for the survey: a 53% response. The average age of responders was 63.9 years (SD 11.2 years) and approximately half were women.

Of the persons who responded to the population survey, 9110 (60.4%) gave their consent to taking part in a follow-up and for their medical records to be reviewed for research purposes. These persons had also reported pain in one of four joint sites (hand, hip, knee or foot) in the previous 12 months when responding to the survey. This was the pool of eligible participants for the individual patient follow-up and analysis in the trial, whose GP medical records were tagged. There were 4702 patient records tagged in the intervention practices and 4408 in the controls.

Active trial participants recruited at the time of consultation about joint pain or OA from this pool of eligible persons numbered 651 in total: 364 in the intervention practices and 287 in the controls. The mean age was 67.3 years (SD 10.5 years) and 59.6% were women. The OA consultation prevalence over the 6 months of the trial recruitment of individuals for follow-up was 651 per 9110 across all practices (7.15 per 100 persons aged ≥ 45 years). The figure was 7.74 per 100 in the intervention practices and 6.51 in the controls (see *Figure 7*).

All 651 eligible patients who consulted with joint pain and OA in the trial period were contacted within 2 weeks of the consultation by postal questionnaire ('post-consultation questionnaire'). There were 525 responses from the 651 eligible patients in all practices (an overall response of 80.6%), 288 from 364 eligible in the intervention practices (a response of 79.1%) and 237 from 287 eligible in controls (a response of 82.5%).

The demographic characteristics of post-consultation questionnaire responders are shown in *Table 16*.

Process outcomes Uptake of NICE core guidance and reported items of self-management (*Tables 17* and *18*).

TABLE 16 Characteristics of eligible patients who have consulted about joint pain or OA and responded to the first post-consultation questionnaire

Participant characteristics	Intervention ($N = 288$)	Control ($N = 237$)
Sex, *n* (%)		
Female	167 (58.0)	146 (61.6)
Male	121 (42.0)	91 (38.4)
Age (years), mean (SD)	66.9 (10.6)	67.7 (10.3)
BMI (kg/m^2)		
Mean (SD)	28.1 (5.1)	28.5 (4.8)
Marital status, *n* (%)		
Married	186 (65.0)	168 (71.0)
Separated	2 (0.7)	4 (1.7)
Divorced	29 (10.1)	13 (15.6)
Widowed	44 (15.4)	37 (15.6)
Cohabiting	10 (3.5)	9 (3.8)
Single	15 (5.2)	6 (2.5)
Employment status, *n* (%)		
Employed	77 (27.2)	59 (25.2)
Not working/retired	206 (72.8)	175 (74.8)
Deprivation Index		
Median (IQR)	21,868 (15,144–28,649)	20,182 (15,989–24,635)
Number of pain sites, *n* (%)		
1	55 (19.1)	45 (19.0)
≥ 2	233 (81.0)	192 (81.0)

IQR, interquartile range.

Parts of this table have been reproduced from Dziedzic *et al.*[168] © 2018 The Authors. Published by Elsevier Ltd on behalf of Osteoarthritis Research Society International. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

TABLE 17 Self-reported receipt or uptake of items indicating quality of care among 525 eligible patients who have consulted about joint pain or OA and returned a post-consultation questionnaire (OA-QIs questionnaire)

Characteristics	Control ($N = 237$)	Intervention ($N = 288$)	
OA QIs (i.e. interventions offered)	*n* (%)	*n* (%)	OR[a] (95% CI); *p*-value
Written or verbal information about joint problem	133 (56.6)	180 (63.2)	2.15 (1.02 to 4.53); $p = 0.045$
Information about treatments	142 (61.5)	186 (66.7)	1.88 (0.85 to 4.13); $p = 0.117$
Advice on self-management of joint problem	128 (54.5)	157 (56.0)	1.22 (0.60 to 2.46); $p = 0.586$
Support on how to help self with joint problem	89 (38.4)	110 (39.0)	1.97 (0.74 to 5.23); $p = 0.174$
Information/advice about exercises, muscle-strengthening or physical activities	101 (43.2)	149 (53.0)	2.18 (1.01 to 4.69); $p = 0.046$
Referral to strengthening or physical activities	61 (26.1)	95 (34.1)	1.75 (0.90 to 3.39); $p = 0.099$
Advice to lose weight[b]	60 (32.1)	66 (30.0)	1.46 (0.47 to 4.60); $p = 0.513$

continued

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 17 Self-reported receipt or uptake of items indicating quality of care among 525 eligible patients who have consulted about joint pain or OA and returned a post-consultation questionnaire (OA-QIs questionnaire) (*continued*)

Characteristics	Control (*N* = 237)	Intervention (*N* = 288)	
OA QIs (i.e. interventions offered)	*n* (%)	*n* (%)	OR[a] (95% CI); *p*-value
Referral to services for losing weight[b]	8 (4.3)	16(7.1)	2.68 (1.09 to 6.60); $p = 0.031$
Need for walking aid assessed	37 (23.3)	53 (29.4)	1.58 (0.83 to 3.00); $p = 0.161$
Need for appliances/aids to daily living	19 (14.8)	13 (10.4)	0.80 (0.31 to 2.07); $p = 0.652$
Paracetamol recommended for pain	119 (53.1)	176 (63.8)	1.50 (0.94 to 2.41); $p = 0.092$
Information about drugs effect provided	124 (52.3)	133 (46.2)	0.45 (0.23 to 0.89); $p = 0.021$
Surgery evaluation	46 (19.4)	50 (17.4)	0.95 (0.21 to 4.24); $p = 0.947$

a Reference is control group.
b Analysis restricted to participants who were overweight.

TABLE 18 Self-reported use of items of care and self-management among 525 eligible patients who have consulted about joint pain or OA and returned a post-consultation (KNEST) questionnaire

Characteristics	Control (*N* = 237)	Intervention (*N* = 288)	
KNEST: medication/treatment prescribed/used	*n* (%)	*n* (%)	OR (95% CI); *p*-value
Paracetamol	165 (69.6)	204 (70.8)	1.05 (0.48 to 2.30); $p = 0.894$
Anti-inflammatory creams/gels (e.g. topical NSAIDs)	159 (67.1)	183 (63.5)	0.60 (0.27 to 1.32); $p = 0.202$
Capsaicin cream	23 (9.7)	23 (8.0)	1.18 (0.46 to 2.99); $p = 0.730$
Anti-inflammatory tablets (e.g. oral NSAIDs)	138 (58.2)	143 (49.7)	0.23 (0.09 to 0.57); $p = 0.001$
Stronger painkillers (e.g. opioids, co-proxamol)	114 (48.1)	130 (45.1)	0.88 (0.35 to 2.19); $p = 0.787$
Glucosamine or chondroitin sulphate	67 (28.3)	90 (31.3)	2.01 (0.70 to 5.80); $p = 0.196$
TENS	29 (12.2)	34 (11.8)	0.88 (0.48 to 1.61); $p = 0.675$
Warmth, heat or cold application	101 (42.6)	137 (47.6)	1.32 (0.58 to 3.01); $p = 0.506$
Walking aids	86 (36.3)	89 (30.9)	0.21 (0.06 to 0.68); $p = 0.009$
Shock-absorbing shoes or insoles	42 (17.7)	52 (18.1)	1.46 (0.50 to 4.30); $p = 0.488$
Appliances and support and braces	55 (23.2)	59 (20.5)	0.70 (0.16 to 2.97); $p = 0.629$
Assistive devices	29 (12.2)	37 (12.8)	1.46 (0.44 to 4.80); $p = 0.488$
Community pharmacy	16 (6.8)	34 (11.8)	2.33 (0.93 to 5.84); $p = 0.071$
Physiotherapy	60 (25.3)	73 (25.3)	0.47 (0.18 to 1.25); $p = 0.131$
Muscle-strengthening exercises	63 (26.6)	117 (40.6)	2.87 (1.37 to 6.02); $p = 0.005$
General fitness exercises	51 (21.5)	67 (23.3)	0.66 (0.25 to 1.75); $p = 0.405$
Diet to lose weight	69 (29.1)	73 (25.3)	0.89 (0.52 to 1.52); $p = 0.669$
Written information about treatments	53 (22.4)	76 (26.4)	1.50 (0.91 to 2.47); $p = 0.110$
Written information about joint problem management	34 (14.4)	71 (24.6)	2.17 (1.25 to 3.74); $p = 0.006$
Written information about OA	38 (16.0)	90 (31.2)	3.29 (1.97 to 5.46); $p \leq 0.001$

TENS, transcutaneous electric nerve stimulation.

Reported receipt of the OA-QIs of care (OA-QIs questionnaire; see *Table 17*) was statistically significantly higher for some items in the intervention practice patients than the control group, notably for provision of written or verbal information about their joint problem, and specifically:

- written information about management of their joint problem and about OA
- information and advice about exercises, muscle-strengthening or physical activities
- among those who were overweight, referral to services for losing weight.

However, receipt of information about the effect of drugs being prescribed was reported by a significantly lower proportion of patients in the intervention arm than controls.

Reported use of items of care for OA (the KNEST questionnaire; see *Table 18*) was also statistically significantly different between patients in the intervention and control groups, notably lower use of oral anti-inflammatory medication and walking aids and higher reported use of muscle-strengthening exercises in the intervention group versus the control group. These differences continued to be observed in the reports of care received or adopted by patients, recalled over an extended period of 3 months in the second follow-up questionnaire (data not shown).

Other process measures of implementation of the MOSAICS model of care

Fidelity of intervention At 3 months after their index consultation, 70 persons in the intervention group (29%) reported having consulted with a practice nurse about their joint problem compared with 26 (13.5%) in the control group.

Patient report on consultation In the post-consultation questionnaire, patients in the intervention practices were also asked to report on the consultation with the GP. Responses were obtained from 273 patients. There were 127 (46.5%) who reported that the GP had asked them what they thought the problem was caused by, 184 (67.4%) who reported that the GP told them what the problem was caused by, 99 (36.3%) who reported that the GP explained what OA is and 229 (83.9%) who reported that the GP gave them the OA guidebook.

Physical activity outcome In *Table 19*, the physical activity scale is shown and includes the follow-up measures. There is a difference at baseline but this becomes exacerbated at 3 and 6 months, with the PASE mean score being statistically significantly lower and declining in the intervention group than in the control group. This is further explored in *Table 20*, which examines the 3- and 6-month differences in each component category of the PASE scale. The differences lie almost exclusively in lower walking scores in the intervention group compared with the control group. IPAQ scores were not statistically significantly different between the groups at any point of follow-up.

Enablement Enablement as an indicator of confidence to self-manage is shown in *Table 19*. Patients in the intervention group reported higher mean enablement scores at 3 months but this had become a statistically non-significant difference at 6 and 12 months.

Clinical outcomes At 6 months there were 239 responders in the intervention practices (83.0% response) in patients from the intervention arm and 185 (78.1% response) in the control arm. There were no statistically significant differences in SF-12 PCS at 6 months (*Table 21*) or in the secondary clinical outcomes (data not shown) between intervention and control groups.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 19 Process measures of implementation of MOSAICS: exercise, walking and physical activity, and enablement

Process measures	Control			Intervention			Mean difference[b]	95% CI[c]	*p*-value[c]
	n[a]	Mean	SD	*n*[a]	Mean	SD			
IPAQ									
Post-consultation	171	3124.79	3829.79	200	2745.52	3285.00			
3 months	157	3305.81	4073.19	182	2378.55	2912.15	–693.31	–1446.77 to 60.15	0.071
6 months	144	2519.33	2786.78	181	2200.90	2967.76	–628.88	–1396.63 to 138.85	0.108
12 months	142	3040.81	3459.51	167	2355.51	2413.99	–594.63	–1396.05 to 206.80	0.146
PASE									
Post-consultation	195	147.49	85.28	237	138.71	75.88			
3 months	176	147.50	86.64	203	123.65	71.95	–22.01	–36.53 to –7.49	0.003
6 months	143	136.24	73.15	190	123.02	68.67	–18.01	–33.33 to –2.68	0.021
12 months	142	148.23	77.90	157	134.22	69.55	–20.61	–36.42 to –4.80	0.011
Patient enablement									
Post-consultation	–	–	–	–	–	–	–		
3 months	202	2.61	3.24	253	2.82	3.16	0.86	0.13 to 1.61	0.022
6 months	178	2.28	2.96	224	3.21	3.43	1.42	–1.50 to 4.35	0.340
12 months	162	2.60	3.19	198	2.80	3.18	2.47	–0.90 to 5.85	0.151

a Number in the sample.
b Calculated as mean difference for intervention – control score by linear mixed modelling adjusted for age, sex, practice size and corresponding baseline measures ('clustering' accounted for in the mixed model).
c CIs and *p*-values relate to adjusted results from the linear mixed modelling.

TABLE 20 PASE components analysis: intervention vs. control

Components of the PASE questionnaire	Mean difference[a]	95% CI	*p*-value
Exercise/strengthening			
3 months	0.02	–0.05 to 0.09	0.584
6 months	–0.02	–0.10 to 0.05	0.550
12 months	0.003	–0.08 to 0.08	0.933
Walking			
3 months	–0.27	–0.48 to –0.06	0.010
6 months	–0.31	–0.52 to –0.09	0.005
12 months	–0.18	–0.41 to 0.05	0.125
Light sport/recreation			
3 months	–0.16	–0.32 to 0.01	0.061
6 months	–0.11	–0.28 to 0.06	0.187
12 months	–0.05	–0.23 to 0.12	0.560

TABLE 20 PASE components analysis: intervention vs. control (*continued*)

Components of the PASE questionnaire	Mean difference[a]	95% CI	*p*-value
Moderate sport/recreation			
3 months	0.07	–0.01 to 0.15	0.077
6 months	0.03	–0.05 to 0.12	0.391
12 months	0.04	–0.05 to 0.12	0.406
Vigorous sport/recreation			
3 months	–0.06	–0.13 to 0.01	0.100
6 months	–0.02	–0.09 to 0.06	0.602
12 months	–0.02	–0.10 to 0.06	0.623

a Calculated as mean difference for intervention – control score by linear mixed modelling adjusted for age, sex, practice size and corresponding baseline measures ('clustering' accounted for in the mixed model).

Note

Higher PASE scores denote greater levels of physical activity. Negative differences in both *Tables 19* and *20* mean that the intervention group had worse scores on that item. The absolute scores for each PASE component are smaller than overall PASE scores (data not shown).

TABLE 21 Primary outcome measure (SF-12 PCS) by study groups including sensitivity analyses

	Control			Intervention					
SF-12 PCS	*n*	Mean	SD	*n*	Mean	SD	Mean difference[a]	95% CI[b]	*p*-value[b]
Post-consultation	231	36.48	11.00	280	36.49	11.48			
3 months	204	38.12	11.58	250	38.03	12.32	–0.31	–1.87 to 1.25	0.698
6 months	180	38.89	12.00	229	38.98	12.12	–0.38	–2.31 to 1.54	0.696
12 months	166	39.22	11.83	200	38.78	12.57	–0.89	–3.33 to 1.55	0.476
Sensitivity analysis of SF-12 PCS at 6 months									
1. In multilevel model, practices replaced with individual GPs	–	–	–	–	–	–	–0.39	–2.31 to 1.54	0.696
2. CACE	–	–	–	–	–	–	–0.11	–4.96 to 4.74	0.963
3. Two-stage cluster level analysis[c]	–	–	–	–	–	–	n/a	0.824 to 0.838[c]	0.832[c]

n/a, not applicable.

a Calculated as mean difference for intervention – control score by linear mixed modelling adjusted for age, sex, practice size and corresponding baseline measures ('clustering' accounted for in the mixed model).

b CIs and *p*-values relate to adjusted results from the linear mixed modelling.

c Two-stage cluster level analysis using permutation test provides test with the advertised Type I error in conjunction with 95% CIs.

In the CACE sensitivity analysis (see *Table 21*), there were no statistically significant differences in the primary outcome between those participants who received the full intervention (GP and practice nurse) and those who did not.

Pre-specified subgroup analysis After stratification by SF-12 PCS scores (< 36.9 vs. ≥ 36.9) at the immediate post-consultation assessment, there were no significant differences in the SF-12 PCS at 6 months in the intervention versus control groups.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Conclusion

Self-reported outcome measures from the subgroup of individuals within the cluster RCT who were recruited to follow-up for the individual patient-level analysis confirmed the practice-level analysis of the 'whole practice' populations presented in study 4.1:

- interventions developed as part of the MOSAICS workstream resulted in improved delivery of some of the items of NICE core guidance to patients aged ≥ 45 years who consulted with joint pain and OA in the intervention practices compared with similar consulters in the control practices, and a modest short-term increase in patients' confidence to self-manage, but
- nurse-led OA follow-up clinics in the intervention practices were only used by a minority of patients, and
- this implementation did not result in an improvement in physical function or pain among patients followed up in the intervention practices compared with those in the control practices.

One possible explanation for the lack of clinical change is the relatively low uptake of the nurse referral clinics in the intervention group. However, there was no evidence from the per-protocol CACE analysis that the group who did see the nurse had better clinical outcomes, although the evidence from the practice-level analysis was that this group had the optimal implementation of quality of care according to NICE guidance.

A second explanation may lie in the data on walking and physical activity. The evidence from the OA-QIs and KNEST questionnaire was that a higher proportion of patients in the intervention group reported receiving advice on muscle strengthening and carrying out muscle-strengthening exercise, but there was no difference in the proportions doing general fitness exercise in the two arms. The PASE scores, surprisingly, indicated that the intervention group had statistically significantly lower scores (i.e. less physical activity) than the control group throughout follow-up, not explained by initial differences between the arms post-consultation. The analysis of PASE domains suggested that the difference lay almost exclusively in lower walking scores in the intervention group, among which there was also a lower reported use of walking stick. The picture painted is of patients informed about muscle-strengthening anaerobic exercise, and carrying it out, but without increasing their walking. This result does not align with the evidence about the benefits of walking and aerobic physical activity (see the BEEP trial) but does suggest an impact of the MOSAICS model of care on exercise patterns.

However, from the patients' perspective, there are components of the NICE guidance and the MOSAICS model of care that are important regardless of their overall effect on pain and disability levels. First, information provision is highly valued by patients and there was clear evidence that this was improved, notably with written material, in the intervention group. Second, patterns of analgesia moved towards safer recommended behaviour (reduced use of oral NSAIDs), as judged by the self-reported process outcomes in this individual-level analysis, supporting the findings from the routinely recorded outcomes in the practice-level analysis. The lowered level of advice about medication in the intervention group may reflect this lower use of NSAIDs. Third, there was some evidence of a shift towards non-pharmacological interventions in the content of the information provided, referral of overweight patients and a short-term increase in patient confidence about self-management.

The search for wider explanations of why the implemented best-evidence care does not improve patient-reported clinical outcomes must consider the appropriateness and timing of outcome measurement, and the possibility that core NICE interventions developed for single-site OA, such as the knee, may be insufficient to shift long-term pain and disability for all patients with OA who are seen in primary care, most of whom have symptoms in multiple joint sites.

However, for some important items, we conclude that the MOSAICS intervention, which provides a model OA consultation with training for GPs and practice nurses, and an OA guidebook appears to add to the effect of the e-template in delivering core NICE recommendations and quality OA care in general practice.

One strength of this study was the initial population survey to identify potential eligible persons before any consultation and recruitment in the trial period. This addressed a major concern for cluster RCTs, namely patient

selection differences between intervention and control practices. Survey responders with joint pain were, on average, older and more likely to be female than those with no joint pain and responders with OA were even more so. This means prevalence estimates are likely to be inflated; however, this response bias is unlikely to have affected estimated associations between baseline characteristics and future events such as treatment uptake.

One potential weakness here, and in study 4.1, was the baseline difference in median rank of area social deprivation between intervention and control practices because intervention practices were, on average, in more deprived localities. However, individual practices were distributed across comparable ranges in intervention and control, and area deprivation scores for individual patients selected for study 4.2 were lower (meaning greater deprivation) in the control group compared with the intervention group. It seems unlikely that this would substantially affect the quality-of-care comparisons.

Study 4.3: cost-effectiveness of the MOSAICS intervention: health economic evaluation

Parts of this section are summaries of material in Oppong *et al.*[172] © The Author(s) 2018. Published by Oxford University Press on behalf of the British Society for Rheumatology. This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com.

Introduction/background

Aim The aim was to determine the cost-effectiveness of the template and model osteoarthritis consultation (MOAC) in comparison with template alone in patients who consult with OA in primary care.

Methods A cost–consequence analysis was conducted, describing all the important results relating to costs and consequences [clinical outcomes, EQ-5D, SF-6D, ICEpop CAPability measure for Adults (ICECAP-A)].

- Costs of the intervention were obtained in discussion with the study co-ordinators. Zero costs were assigned to the control arm.
- Costs associated with over-the-counter medications were obtained from patient self-report in the questionnaires or by allocating costs to the drugs mentioned in self-report.
- Resource use was valued by obtaining unit costs (2012/13 prices) from standard sources such as Personal Social Services Research Unit (PSSRU),[173] BNF[174] and NHS reference costs[175] and applying them to resource use data from the trial data sets.

An incremental cost–utility analysis was then undertaken using patient responses to the three-level EQ-5D questionnaire (at 'baseline', 3, 6 and 12 months) to estimate incremental costs, QALYs and net benefits from a UK NHS perspective. Uncertainty was explored through cost-effectiveness acceptability curves (CEACs). Sensitivity analysis explored the robustness of the results.

Full details of the data and methods, including information on resource use and associated costs, used to develop the economic model for delivering the MOSAICS intervention, are included in *Appendix 2* (see *MOSAICS Health Economics Methods*).

Results A total of 525 participants from eight general practices were included. There were no statistically significant differences in health outcomes and capability between the intervention and control arms at any time point. Occurrence of visits to the orthopaedic surgeon was lower in the intervention arm versus control arm (mean number of visits per person over the designated follow-up period: 0.28 vs. 0.53; $p = 0.02$). There were no other statistically significant differences in primary or secondary care resource use between the two arms of the cluster trial over the 12 months' follow-up.

Cost–utility analysis shows that the intervention was associated with an incremental cost of –£2.78 (95% CI –£67.61 to £62.03), an incremental QALY of –0.003 (95% CI –0.03 to 0.02) and a 40% chance of being cost-effective at a threshold of £20,000 per QALY gained.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

The percentage of participants who took time off and the associated productivity costs were both lower in the intervention arm.

Conclusion Implementing NICE OA guidelines in primary care using the MOSAICS intervention does not lead to increased costs and appears to reduce demand for orthopaedic surgery and time lost from work; however, QoL was little reduced.

Group 5 studies: qualitative evaluation of the MOSAICS studies

Parts of this section have been reproduced from Morden *et al.*[176] © 2014 Morden *et al.*; licensee BioMed Central. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly credited. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated; and Morden *et al.*[177] © Morden, Brooks, Jinks, Porcheret, Ong and Dziedzic. 2015. Published by Emerald Group Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 3.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence may be seen at http://creativecommons.org/licences/by/3.0/legalcode.

Study 5.1: implementing a complex intervention in practice – an evaluation using a theoretical framework

Introduction/background

One approach to the many challenges of implementing results of research into routine practice is to consider implementation from the start of a research project by using a theoretical framework. Doing so can help to ensure research (in this case the MOSAICS model of care) is more relevant to the NHS and has greater potential to benefit patients.[137] Such an approach can highlight factors that may influence the likely success, outcome or long-term implementation of an intervention.

Aim

To understand the way in which health professionals (GPs and practice nurses) made sense of the new intervention, how they operationalised it and what influenced their continued implementation or rejection of the new approach.

Methods

Several theoretical models have been developed to understand the social processes associated with the uptake of complex interventions. However, they pay limited attention to the early stages of 'sense-making' work that people do when faced with how to put a new intervention into practice, or the work on relationships that people do in order to sustain a 'community of practice' around a new intervention. In order to remedy this, we adopted the NPT[137] to frame the MOSAICS studies and analyse the data. A detailed description of the conceptual underpinning of our approach and its application in the MOSAICS studies and more widely in primary care is provided in Ong *et al.*[178]

A qualitative design was utilised. In order to obtain a variety of perspectives of the same phenomena,[179] two types of data collection strategies were used. First, there was direct observation. Second, group and individual interviews were conducted with GPs and nurses who participated in the MOSAICS studies, as well as staff from the NIHR Primary Care Research Network (PCRN) who acted as intermediaries between the research team and practices. Observations and interviews were conducted at multiple time points.

Observations included meetings introducing the complex intervention to the practices (all eight participating practices were observed); the four intervention practices were also observed in subsequent meetings and training sessions and during post-study feedback meetings. All practice nurses ($n = 9$) were

observed delivering their clinics (a total of 27 clinics across all intervention group practices were observed). Interviews were conducted with network staff at the introductory stages of the intervention ($n = 5$). Samples of GPs ($n = 10$) and nurses ($n = 5$) were interviewed after receiving training and nine GPs and four nurses were interviewed post intervention.

More details of the conceptual underpinning of this work and of the methodological processes of data collection and analysis have been published by Morden *et al.*[180] and Ong *et al.*[181]

Results

'Sense-making' GPs varied in their views of the relevance of the new intervention to their everyday clinical work and their organisational priorities. Although they recognised that OA was common and problematic for patients, and that they lacked up-to-date knowledge and skills to treat and manage the condition, OA did not appear to be a priority for them because many GPs perceived it to be related to ageing and it did not figure as a national policy and payment priority. The new intervention was viewed as potentially helpful in systematising OA care (by use of the e-template) and because of the opportunity it provided to refer to the practice nurse for further support and advice. The GPs also considered the guidebook a valuable aid for patients.

The practice nurses saw the nurse-led clinic more positively as an extension of their skill set because it allowed them to explain OA to patients, discuss treatment options, support self-management and take account of psychosocial factors.

The research team highlighted the need to establish relationships, understand practice dynamics and engage key decision-makers as being central to the process of engaging health-care professionals with the study.[178]

Factors relevant to participation in the MOSAICS studies To facilitate recruitment and operationalisation of complex interventions, the research team needed to pay attention to the existing clinical concerns, internal power dynamics and workloads within individual primary care organisations. Logistics and existing communication channels influenced how health-care professionals made sense of the study and organised their roles in relation to it.

Most GPs saw the adoption of the template in their decision-making about OA care as useful and the practice nurses were prepared to take on the bulk of the work required for the new intervention. They appeared to consider this division of labour appropriate to their role which made the study acceptable.

'Routinization' of new complex interventions in practice Practices did not think that they would be able to routinely use the intervention as designed beyond the period of the funded trial. Continuing with OA-specific nurse-led clinics was resource intensive and other incentivised priorities took precedence. However, clinicians suggested that they would be able to continue with selected components of the intervention and embed them within routine practice, namely using the communications skills and technical disease-specific knowledge in routine consultations.

Conclusion

Using a theory-based approach such as NPT can help researchers to develop and introduce new interventions to primary care settings and ensure uptake. The interplay of system-, group- and individual-level factors shapes the specific context of primary care. The effects on the uptake of new interventions vary depending on factors such as timing (e.g. a change in policy), actors involved and practice dynamics, perceptions of own professional identity, and patient needs and demands.

The manner of introduction appears to be highly relevant, with levels of flexibility and negotiation being crucially important. Furthermore, it is essential to understand the dynamic process of adaptation as an integral part of implementation and routinisation, and to assess its contribution to eventual outcomes of a

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

study and long-term implementation. In the case of MOSAICS, selective adaptation of elements of the intervention was the route to continued use in practice.

Study 5.2: experiences of model consultations – patients

Introduction/background

Literature[23] and research findings which fed into the MOSAICS studies suggested that patients, when they consult GPs, are not optimally treated for OA, often not given a diagnosis, prognosis or information, and often not provided with support for self-management. The intervention in our study is a new way of offering OA care and an important ambition was to understand patient perspectives of the new approach, what worked, what went less well and why.

Aim

The aim of this study was to explore and understand what patients thought about the intervention, if it improved their experience of OA care and if it encouraged the uptake of the core NICE OA treatments (self-management).

Methods

In-depth interviews were used because they can yield rich sources of data on people's experiences, opinions, aspirations and feelings.[182] They enable respondents to tell their own stories in their own words and the meaning that people attach to events can be revealed.[183]

A convenience sample was used and patients who had consulted at intervention practices were recruited. Weekly medical record downloads from intervention practices identified patients who had consulted for OA. These patients were issued baseline and 3 month 'post-consultation questionnaires' as part of the broader study evaluation. From returned questionnaires, patients were identified who indicated they had seen both the GP and practice nurse. Potential participants were sent an invitation letter and information sheet offering them the opportunity to take part in this interview study. A total of 29 patients who had consulted for OA and had received the MOSAICS intervention volunteered to take part.

All interviews were audio-recorded and professionally transcribed verbatim. Thematic analysis was undertaken using some of the principles laid out by Grounded Theory, in particular focusing on identifying emergent codes, developing themes and constantly comparing data and coding.[184]

Details of the methods of data collection and analysis have been published in Morden *et al.*[180]

Results

Findings from this study suggest that all patients gained something from the intervention, but that separate components of the intervention may not have been universally helpful for all patients. There are three key areas that participants discussed in relation to the new intervention: (1) information and explanations about their condition provided during the intervention, (2) perspectives on holistic care and support received and (3) responses to self-management advice and support, in particular, exercise/activity.

1. Information and explanations
   All participants compared the intervention with previous experiences of consulting for OA. Participants discussed frustrations with previous consultations stemming from a lack of clear diagnosis, the perception that they had received few treatment options and answers, and time-consuming referrals to secondary care (which also provided few answers). Participants suggested that consulting a GP for OA as part of the new intervention featured less use of the term 'wear and tear' to describe joint pain and resulted in provision of advice about exercise, provision of the OA guidebook and referral to see a nurse. The guidebook was deemed useful because of the explanation it provided regarding what causes OA, how OA affects people and a prognosis. Equally, being able to discuss OA with a nurse, clarify concerns and receive personalised information was appreciated by participants.

2. Holistic care
   The theme of holistic care and support featured differences by sex. Female participants described the value of time and attention given to discussing personal worries and emotional issues relating to pain. They also felt that, in combination with information provided by someone perceived to be a clinical expert, the time and attention provided by nurses offered a sense of 'legitimation' because previously they did not think they had been taken seriously by clinicians and their social contacts. Female participants suggested that, in combination with the information about OA, they were better able to 'cope' and had a better sense of 'well-being'. Men discussed how information contained in the guidebook relating to emotions, thoughts and feelings was useful because it offered reassurance that what they were experiencing was normal. However, male participants did not recount benefits gained from discussing personal worries, gaining legitimation or a better sense of well-being and being able to cope as a result of consulting with nurses. They focused their accounts around obtaining technical information about OA and a 'cure' (which was often interpreted as not forthcoming).
3. Uptake of advice about exercise
   Uptake of advice about exercise, based on the qualitative evidence, was variable. Patients understood the rationale for doing aerobic and muscle-strengthening exercises and reported trying them, but they did not necessarily continue with them. One reason centred on participants' 'sense-making' regarding the effectiveness of exercise. Put simply, participants monitored and observed symptoms (pain, swelling and function) for signs of improvement in the short term and mid-term, and their continued use of exercise (in particular, muscle strengthening) depended on 'proof' of effectiveness. A second factor was the presence (or absence) of comorbidities which influenced people's ability to exercise (particularly aerobic exercise). Finally, participants drew attention to the notion of appropriate places to exercise, with the home and immediate neighbourhood considered to be 'off limits' for exercise. Thus, the availability of affordable venues that felt comfortable (i.e. age appropriate, not feeling self-conscious because of body size) was another factor influencing patient's engagement with exercise as recommended by the nurse.

Conclusion

The intervention overcame some of the limitations and challenges posed to patients by usual care, namely by providing a diagnosis and information about OA as a disease entity. To differing degrees, the intervention offered holistic support to patients, more pronounced for female patients. However, uptake of advice about physical activity and exercise was not universal and was dependent on environmental and individual factors that the intervention, from patients' accounts, did not or could not address.

Study 5.3: experiences of model consultations – general practitioners

Introduction/background

The MOSAICS studies intervention was a new way of offering OA care, which involved GP training programmes and potentially added work for GPs in consultation and recording. It was important to understand GP perceptions about using this new approach to OA care and about how useful the intervention was for patients.

Aim

The aim was to explore GPs' perspectives about delivering the MOAC intervention and about its strengths, weaknesses, acceptability and feasibility.

Methods

GPs ($n = 9$) from intervention practices volunteered to take part in semistructured telephone interviews. Post-intervention feedback meetings were observed by two members of the research team (AM and BNO). Observation as a qualitative research method involves the researcher 'going into the field' and describing and analysing what has been seen, what people do and what people say, therefore, illuminating behaviour and interactions in natural settings[179] and aims to identify the meaning of, for example, a novel

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

intervention to people in that setting.[185] All data were analysed using on-going inductive coding, subsequent theme development and constant comparison.[184]

Results

The GPs focused on the gains in medical knowledge from participating in the study, stating that knowing more about OA and its management made it feel like less of a 'heart sink' consultation. They felt able to offer more treatment options, provide better information to patients and, therefore, thought they gave better care. As a result, they felt more 'positive' about seeing patients with joint pain.

The GPs thought that the OA guidebook and the new style of consultations provided a means of 'empowering' patients to take responsibility for their condition and alter a perceived relationship of patient dependency on the GP. They also thought the OA guidebook was a useful tool to help back-up key messages about the importance of exercising and taking prescribed medications. GPs thought that the format of the intervention was useful because, by being able to provide a guidebook and/or refer to a nurse, they could 'close off' consultations in a way that felt comfortable; the GPs felt that they were giving the patient tangible support, while not having to increase their own consultation time. However, GPs did express some concerns that referral to the dedicated nurse-led OA clinic might add to the burden on patients with multiple morbidities. Some also doubted that the nurse referral would be suitable for all patients. Examples of poor candidates for referral that GPs mentioned were patients lacking motivation and patients with a particular agenda, such as wanting surgery.

Conclusion

In summary, the intervention was deemed to be positive because it 'up-skilled' clinical knowledge and provided additional treatment and management strategies for patients with OA, and it fitted with GPs' existing ways of working, thus requiring minimal change in their routines. The fact that GPs felt that they had more to offer to OA patients benefited their individual practice and relationship with patients, and reflected greater confidence in using and implementing NICE guidelines.

Study 5.4: experiences of the model consultation – nurses

Introduction/background

Practice nurses potentially play a key role in the delivery of supported self-management for musculoskeletal conditions[186] and they were an important part of the delivery of optimal OA care in the MOSAICSs framework. Because the intervention that was under trial was a new way of offering OA care, qualitative research was undertaken to explore and understand the nurses' experience and perspective.

Aim

The aims were to:

- investigate how nurses delivered the intervention and to understand what happened in practice during the practice nurse OA consultations
- explore nurses' experiences of delivering the intervention
- investigate nurses' perspectives about the strengths and weaknesses of the new intervention.

Methods

Two separate methodological approaches were adopted.

1. Nurses ($n = 4$) from across the intervention practices who participated in the delivery of the trial volunteered to take part in semistructured interviews after the intervention had ended. Interviews were used because they can yield rich sources of data on people's experiences, opinions, aspirations and feelings.[182] They enable the respondent to tell their own stories in their own words and the meaning that people attach to events in social settings can be revealed.[183]

2. All nurses in the intervention practices ($n = 9$) were observed delivering intervention clinics. A total of 27 clinics across all practices were observed. Post-intervention feedback meetings were observed by two members of the research team (AM and BNO). Observation as a qualitative research method involves the researcher 'going into the field' and describing and analysing what has been seen, what people do and what people say, therefore, illuminating behaviour and interactions in natural settings[179] and aims to identify meaning for people in that setting.[185]

All data were analysed using on-going inductive coding, subsequent theme development and constant comparison.[184]

Results

The nurses enjoyed delivering the intervention and, in particular, focused on the gains in medical knowledge about OA that they attained from participating. They made a distinction between the sense of achievement that they gained from 'good' and 'motivated' patients who responded to their support and advice and patients who they thought were not motivated and whom they could not help.

Nurses suggested that, for the 'motivated' patients, the intervention was beneficial because it helped patients to self-manage, in particular with regard to pain and maintaining valued activities which improved health outcomes and QoL. However, they suggested that 'demotivated' patients were more of a challenge and nurses were unsure if they could help them.

Nurses thought the OA guidebook was a useful tool to help back up some of the key messages in their consultations, in particular generic advice about staying positive, the importance of exercise/activity, weight loss and explaining OA as a disease. But, the nurses implied that it could have benefited from containing details of muscle-strengthening exercises because they had to rely on other sources such as Arthritis Research UK (Chesterfield, UK) information sheets to promote muscle-strengthening exercises.

With regard to implementing a whole systems approach (WISE) in the consultations, it seemed, following the observations, that the nurses structured the consultations around promoting exercise and lifestyle changes, which was sometimes to the detriment of psychosocial issues (e.g. work or feelings of stress). This may be because the training focused primarily on the implementation of the NICE OA guidelines and the nurses considered that shaping the consultation following the core intervention was a priority.

Observations also revealed that the nurses struggled to engage in discussions with patients about barriers to exercise, underutilised effective motivational techniques and tended to reinforce core messages about the benefits of exercise in an effort to deliver patient education. Thus, trained practice nurses were challenged in providing support to patients who had difficulties engaging with their advice to exercise. Nurses tended to conceptualise self-management as meaning patient adherence to their particular recommendations, goals or targets. Although overall style of consultation and level of patient participation varied, most consultations tended to be nurse directed with a focus on promoting specific exercises identified in the guidelines (e.g. strengthening exercises). They also tended to follow a scripted approach, which might overlook patient needs and concerns in relation to general physical activity and focus on nurse-led goal-setting. This, in particular, may have resulted in the counter-intuitive finding from the quantitative data that walking declined in the group seen by nurses while strengthening exercises increased.

Conclusion

Nurses valued the intervention and found it worthwhile in terms of their individual practice and professional development. However, some tensions emerged between how nurses positioned the utility and success of the intervention (ability to motivate patients) and the way that they engaged in discussions with patients. Although nurses interpreted the intervention as a success from their perspective of delivering the core interventions and suggested that patient characteristics may contribute to 'failure' in uptake of self-management, challenges with communication and operationalising behaviour change and communication strategies also played a role. Additional feedback and mentoring as support for nurses in delivering the model OA consultation intervention might address some of these issues.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Group 6 studies: wider dissemination of the MOSAICS interventions

The original idea of 'learning sets' as the basis for the dissemination phase of MOSAICS was laid out in the MOSAICS protocol. It focused on direct engagement of primary care trust (PCT) managers. This objective became impossible to develop and implement because of the major reorganisations in the NHS that were occurring at the time. In particular, it became clear that setting up 'learning sets' during the period when new structures such as Clinical Commissioning Groups (CCGs) were forming would not be possible. An alternative strategy was adopted to meet these new circumstances.

First, a 1-day event was held at the end of the study to present interim findings from the MOSAICS studies to all those who had participated, to the wider patient groups and to local health professional leaders. This provided further proposals for improving the intervention, the way in which it should be introduced and supported in primary care and how to maintain it in practices that had adopted the approach. For example, there was support for the potential to continue to use the computer template as a way to underpin a systematic approach to OA consultations and as a means to introduce GPs who were newly arriving at participating practices to the NICE OA guidelines.

Second, two approaches were adopted to achieve wider dissemination: (1) development and evaluation of modifications to the training programme and (2) extension of the intervention to a new set of practices beyond the MOSAICS participants. These are described below.

Study 6.1: revising the training programme – engaging control practices post trial

Introduction/background

The research team reviewed the training programme after analysing the feedback from the intervention practices and designed a new 4-day programme for nurses and 3-day programme for GPs. The training was made available to staff from the four control practices. The control practices received this refined training package after the end of the MOSAICS trial evaluation period.

Aim

The aim was to establish how the revised training package could prompt and enable changes in practice with the aim of enhancing care for patients with OA.

Methods

The method adopted was a single-event facilitated group discussion.

With an ethical amendment and informed consent, all health-care professionals in the primary care teams in the four control practices who had taken part in the post-trial training (17 GPs and 6 nurses) were invited to attend a final session, which formed the basis for the facilitated group discussion.

The discussion was led by a clinical researcher (rheumatologist), with a semistructured schedule that encouraged the group to think about how they could use the skills they had learned, in addition to capturing their perceptions about the training. All discussions were digitally recorded and transcribed, and thematic analysis was used.

Results

Three practices participated in the facilitated group discussion, with one declining further involvement. Participants described individual changes they had made to their practice since the training. However, the process of now having dedicated time scheduled for group discussion between nurses and doctors enabled plans to be made for practice-level changes such as referral pathways for OA patients.

The participants discussed areas of need in primary care management of OA, particularly around the use of exercise, barriers to which were compounded by perceived lack of existing resources. The 'whole practice' approach to training, opportunity for individual reflection between sessions and the research-driven content were elements that were perceived as particularly valuable.

There were two unexpected findings. First was the way in which primary care teams felt that health-care support workers and possibly other team members, such as receptionists, may be best placed to deliver and monitor interventions for OA. Second was that GPs and practice nurses reported benefits to their management of other long-term conditions as a result of the training. They also perceived that using nurse time to treat OA would not negatively impact on practice resources as it would benefit the practice in other ways (e.g. better weight management or blood pressure control).

Conclusion

The training had been well received, had unique aspects that were perceived as addressing areas of real need in primary care and had surprise potential benefits for management of other long-term conditions. Further roll-out of the intervention may usefully consider offering training to further members of the primary care team and include timetabled facilitated discussion for participants to generate locally relevant changes in practice from the skills learned.

Study 6.2: extending the MOSAICS intervention to new practices via the new NHS

Introduction/background

A local CCG was engaged through some of its practices that had participated in the MOSAICS studies. After a combination of grass roots interest and support generated by the training, and with GPs and practice nurses reporting greater confidence in managing OA, 15 other practices in the locality covered by the CCG now wished to implement the NICE OA guidelines using refined electronic tools, training and patient information, and using local expertise to champion the spread of innovation.

Importantly, the CCG (the South Shropshire CCG) recognised that the MOSAICS approach would help to address some of the challenges in its plans and the priorities for improving quality in primary care. The CCG worked with the MOSAICS research and user groups, with support from NICE and Arthritis Research UK, to submit and secure a Regional Innovation Fund award from NHS England to introduce the MOSAICS interventions across all its practices. Details of this new work [Joint Innovation of Guidelines for Osteoarthritis in the West Midlands (JIGSAW)] are outlined below.

Aims

The overall aim is to support primary care in addressing the unmet needs of adults consulting for OA through systematic implementation of NICE OA guidelines at practice level. The specific objective of the JIGSAW project is to refine and implement MOSAICS innovations as a pilot excellence pathway for OA in primary care.

The information learned from the MOSAICS studies will be utilised to:

1. embed a template in GP clinical records systems to prompt the recording of quality indicators of OA care
2. provide high-quality patient information – the OA guidebook (www.keele.ac.uk/media/keeleuniversity/ri/primarycare/pdfs/OA_Guidebook.pdf)[163] – developed by patients for patients and revised according to the NICE update in 2014[153]
3. adopt a model OA consultation with GPs and practice nurses that is transferable to other long-term conditions
4. develop a ‘training the trainers’ model to offer:

    i. practice-based training in OA, with one session for the whole practice followed by two brief, intensive sessions for the GPs, including a knowledge update and skills-based training involving simulated patient consultations in a group setting.
    ii. a 2-day programme of training for practice nurses including knowledge components and skills-based sessions on advising and demonstrating exercises, facilitating support for self-management and advising on medication.

5. develop a patient completed audit tool.

© Queen’s Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Methods

The JIGSAW has identified 15 general practices in South Shropshire as pilot sites. The CCG has been engaged as well as the individual practices. The MOSAICS research team and user group are collaborating in this project, but there are methodological innovations that will develop the dissemination beyond the MOSAICS approach, building on lessons learned from the MOSAICS studies, notably:

- 'training the trainers' as the basis for local delivery of training programmes for GPs and practices nurses in the OA consultation
- recruitment of four GP and two practice nurse local clinical champions as the basis for leading on the incorporation of electronic OA templates into practice and to support colleagues in clusters of up to four practices (workshops at two practice-based meetings introduced them to their roles)
- a 2-hour evening meeting introduced JIGSAW to all practices and practice staff, including clinical champions, and provided patient stories, the evidence from MOSAICS and demonstration of the e-templates
- a dedicated Health Informatics Specialist from Keele supported local practices to install the template. Training digital versatile discs (DVDs) were produced
- the template formed the basis for continuing the audit of the uptake of core NICE OA recommendations.

Results

Out of the 15 practices invited to the JIGSAW launch, 12 were represented. Champions had varying degrees of success in engaging practices with OA updates, but 11 practice staff (practice nurses and support workers) attended a 2-day training package to enhance consultation skills for management of OA and other long-term conditions. E-templates were installed in 13 practices and patient guidebooks for OA used in consultations in these practices.

Additional CCG funding was then provided to allow and support practice nurse time for OA consultations in general practice. At this point, JIGSAW was adopted by the West Midlands Academic Health Sciences Network (AHSN), with two further CCGs (Telford and Wrekin, and North Staffordshire) recruited to take part. The South Shropshire CCG then appointed a project manager to support the adoption of JIGSAW across all 44 practices in the CCG. Finally, NICE formally endorsed the OA e-template as a tool for implementing their 2015 OA quality standards.[153]

Measurement of process and patient outcomes from the JIGSAW project lies beyond the remit and timing of this NIHR programme, and the study to date represents an end point of practical supported implementation and dissemination. The outcome of the innovation will be evaluated by:

- change in clinical variation in management of OA
- practice-based process measures, such as reduction in radiography use and referrals to orthopaedic surgery
- uptake of core NICE interventions captured via the electronic template
- audits of patient-reported quality of OA care, satisfaction with care and access to written information (OA guidebook)
- increasing the number of GPs and practice nurses trained, as well as requesting feedback regarding the training.

Conclusions

The JIGSAW has demonstrated to date that innovations developed and tested in MOSAICS, cocreated by GPs, practice nurses, patients and researchers, can be successfully modified and adapted to deliver an excellence pathway for OA in primary care. Further adoption will be supported by the AHSN, with a pan-European partnership proposed in collaboration with NICE.

Overall conclusions

The main result of the cluster RCT is that uptake of optimal primary care for OA that is aimed at implementing the core treatments recommended by NICE can be increased by training in, and delivery of,

a model OA consultation by GPs and practice nurses over a 6-month period, with no additional costs, reduced visits to orthopaedic specialists and reduced time off work. However, the intervention does not improve patients' pain and physical function over the same period.

The interpretation of this finding should be set in the context of the whole workstream. Individual studies have established that the delivery of NICE-recommended core treatments may be improved by the model GP OA consultation, which is linked to the provision of a resource developed and valued by patients (the guidebook) and supplemented by referral to a practice nurse OA clinic. Introduction of a systematic computerised quality-of-care template, linked to consultation, was acceptable to the GPs and resulted in improved recording of quality of care. In general, the GPs, nurses and patients were positive about this model of enhanced primary care for OA, which was noted in particular by the qualitative component of the mixed-methods suite of studies.

The health economic analysis confirmed the main result of no overall benefit to the intervention in terms of health outcomes. However, the analysis did find that there were no cost consequences for the intervention.

These findings raise major issues for practice, policy and research. OA is the leading cause of disability in older people and there is general agreement that the main push of preventing pain and disability will be situated in primary care, unless major public health shifts in physical activity and obesity occur soon. The assumption that drove this workstream was that the components of the NICE core guidance had sufficient evidence of efficacy in relieving pain and reducing disability, which meant that the main challenge lay in the implementation of these interventions and treatments. The MOSAICS studies have provided evidence about how to overcome this challenge and to achieve implementation of the core guidance, including support for self-management and referral on to other services, such as physiotherapy, if the core interventions are ineffective. However, despite this, the individual-level analysis did not provide evidence of a subsequent improvement in the primary clinical outcome.

Limitations of implementation

Implementation of a complex intervention designed to provide an integrated primary care service for patients with OA that was focused on model GP and practice nurse consultations did prove a challenge. Notably, only a minority of patients saw the nurse and so not all eligible patients were offered exercise, which is known to be effective in improving pain and increasing function in knee OA.[122] In addition, those that did see the nurse appeared not to have adopted the optimal change in exercise. The qualitative studies provided further insight into this to show that the patients understood the importance of exercise, but that barriers persisted for aerobic walking exercise because of comorbidities and lack of a suitable environment to encourage exercise.

It is possible that, despite the success of the training programme for GPs in changing some of their behaviours (as shown by changes in the content of their consultations with simulated patients), there has been less of an impact in practice and the impact has been insufficient to influence patients. However, the patient guidebooks, which met criteria for strong patient engagement in guidebook development and incorporated local information, were successfully delivered as part of the model OA consultation and, in the qualitative studies, patients reported relevant GP behaviours in the intervention arm and evaluated the guidebook as helpful. The e-template study provided evidence that prescribing did change and that relevant advice was recorded as being given. However, there were limitations. Only one-third of patients attended the nurse follow-up clinic in the intervention arm. There may be a number of reasons for this, including explanations for low referrals such as:

- non-referred patients did not have OA (this would happen if the GP fired the template, then considered later on in the consultation that the patient's problem was not due to OA and decided not to refer the patient to the nurse)
- some GPs were untrained because of staff turnover rates
- behavioural change in GP competencies did not translate into changed performance.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

The low referral rate to nurse-led clinics may be one explanation for the lack of clinical improvement. It meant that the full integrated MOSAICS model of care, incorporating improved consultations with the GP and referral to practice nurse-led OA clinics were not delivered as intended. Evidence has been provided that seeing the nurse was linked to full implementation of the MOSAICS method of delivering NICE core guidance, including the provision of written information and advice, and opportunities to discuss exercise and physiotherapy referral. This appeared sufficient to explain differences between control and intervention practices in meeting quality-of-care criteria for provision of written information.

However, the fact that only a minority of patients saw the practice nurse cannot finally explain the lack of improvement in pain and disability outcomes in the trial, as even in the subgroup who saw the nurse there was no difference between the intervention and control groups in these outcomes, although it may have affected power to show a difference. There was clear evidence in the practice-level analysis that the combination of GP consultation and nurse referral was achieving the highest level of implementation of NICE core guidance and quality of care. The qualitative studies confirmed the nurses' enthusiasm for the new model of care and provided evidence of patient engagement being driven by the nurse-led clinics, but they also raised some questions about the clinics, based on patients' judgement of the nurse intervention and from direct observation and audio-recording of the nurse-led clinics. In addition, the GPs' use of the template did reduce with time and a similar drift may have occurred in the content of their consultations. Maintenance of behaviour change in clinicians and patients is a challenge for the primary care of OA patients.

Limitations of the intervention

It is possible that crucial elements of an effective intervention were lacking. It is notable that OA treatment remains focused on single-site disease and yet the evidence from this workstream emphasises that it is the person with multiple joint pain who is most commonly seeking help. Elements of primary care may still not be properly targeted or shaped for the majority of patients with multiple joint pain. The emphasis in the NICE guidelines on the importance of advice about exercise and weight loss may be an insufficient intervention in primary care; the guidelines recommend a whole range of further adjunctive treatments should the core intervention be ineffective.[15] The evidence about effectiveness of physiotherapy-led exercise interventions in reducing pain and disability in the short term among OA patients (reviewed and emphasised in *Chapter 4, Introduction*, describing workstream 3 of the programme) may mean current GP or nurse advice alone is insufficient to improve activity to levels which will relieve pain and disability. The finding about the shift in balance towards muscle strengthening and away from aerobic exercise in the intervention practices supports this possibility. Increasing the availability and use of physiotherapist-led exercise and activity as part of a revised MOSAICS intervention is one potential way forward.

However, this raises issues of resource. One question for future research is whether or not there are subgroups of patients in primary care for whom the NICE core interventions are both sufficient and effective. One important issue is the context of multimorbidity, as the baseline population survey confirmed that this is the most frequent state among older people in the community reporting joint pain. There is a continuing need to investigate the role of multimorbidity in determining differential outcomes in primary care patients with OA and how best to manage this.

Limitations of expectation

It is important to consider if process measures of successful delivery of care might be valued as important in their own right, given that there is evidence from elsewhere that the treatments and interventions are worth doing and that feedback from patients and clinicians indicate that the interventions produce a clearer and more positive model of care for patients with OA. Therefore, a second question for future research arising from this workstream is how far process measures of successful implementation of treatments of known efficacy should be used to drive policy in the absence of strong evidence that they are shifting clinical outcomes, and in the presence of evidence that some patient goals (better information, satisfaction with care) are being met by the model of care. One important issue when considering, for example, the application of evidence, to everyday primary care, that physiotherapy-led exercise and activity is efficacious and effective in reducing pain and disability, is that the evidence may have been gained

outside the context in which it would be applied. One of the arguments for complex trials is that the evidence of efficacy is being examined in its practical context.

Limitations of resource

It must also be accepted that support for self-management provided in the GP consultation and the nurse-led clinic may be limited in its potential for impact on long-term clinical outcomes by the limited resources available to the patient outside the primary care setting. The qualitative study of patients highlighted challenges in the outside world, such as limited access to places to exercise, for patients continuing on a self-management path. Such external barriers will undermine the immediately beneficial effects of advice and information also observed in MOSAICS (e.g. shifts in knowledge, initial engagement in exercise). Similar challenges were highlighted in a German study of why GPs felt that implementation of guidelines for back pain would fail – lack of an external environment to meet the requirements of the guideline.[187] Future research could investigate integration between primary care, public health and community resources to support sustainability of OA self-management.

The last two points lead to the question of how implementation of the MOSAICS model might now proceed. A local CCG has engaged in the first stage of planned dissemination and views the evidence of effective processes of GP training, quality templates, nurse-led clinics and patient-generated information (the 'process outcomes' of the MOSAICS studies) as sufficient basis for the new approach. The finding of cost neutrality provides further reassurance for such implementation. Delivering high-quality care with a reduction in orthopaedic visits and use of imaging are desirable outcomes for health-care commissioners. The NPT model, which has framed the work on implementation throughout this workstream, highlights that there remain barriers to implementation, including clinician variation and lack of resource to deal with the dominant problems of low physical activity and obesity.

The lack of evidence demonstrating clinical effectiveness of the current model of optimising care is an important issue to tackle, as clinicians and patients will need evidence that new models of care do ultimately benefit patients if resources to implement them fully are available. However, the evidence that implementation of the model has achieved a higher quality of care, as judged by current standards, is relevant to policy decisions at the same time as future research tackles some of the questions about how those current standards might need to change in order to improve clinical outcomes.

It follows from this that the alternative explanation is that the MOSAICS model of care was not a sufficient intervention to shift pain and disability, and that its focus on core interventions limited the capacity of the individual-level analysis to assess the effect of the full spectrum of adjunctive treatments recommended by NICE if core interventions fail to benefit. Successful implementation of NICE core guidance itself, as demonstrated by both practice- and individual-level analysis of the MOSAICS model of care, may not lessen pain and disability. Optimal and accessible information for patients may be desirable for many reasons but may not, in itself, result in lessened pain and disability and, using the GP and practice nurse consultation as the means to deliver exercise and weight loss advice (although, again, desirable as a first step), may not deliver the full 'dose' of support and intervention needed to shift patient behaviour sufficiently to achieve the reductions in pain and disability that efficacy trials of, for example, physiotherapy-led exercise and physical activity have demonstrated in patients with OA.

One further question that arises is the conceptual one of the extent to which NICE guidelines will be effective when put into practice. One clear novelty of our study is that it has attempted to test the clinical effectiveness and cost-effectiveness of guideline implementation in practice. We have provided process evidence that health-care professionals and health-care systems can change sufficiently to implement components of guidelines, but patient benefits remain elusive to demonstrate.

Summary: the MOSAICS research story

The aim of the project was to develop and evaluate ways to put NICE core guidance for OA into primary care practice in a combination of implementation study (evaluating ways to put NICE guidance into practice

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

on the assumption that evidence-based clinical efficacy for patients would follow) and effectiveness study (actually investigating whether or not clinical efficacy does follow in real-world practice).

The implementation study

- This focused on improved resources/technologies for primary care (including a patient-developed guidebook, computerised consultation template, nurse-led OA clinic, training for GPs and nurses) designed to promote behavioural change in GPs and nurses, and (via guidebook and advice) self-management behaviours in patients.
- The MOSAICS intervention demonstrated that a 'whole-system' implementation of NICE guidance can be achieved in primary care and can change clinician behaviour and support patient self-management. The study produced evidence of improved capacity for NICE implementation at the:
    - GP level (the training delivered behavioural change, guidebooks were handed out to patients, items on the template were carried out, prescribing changed and nurse referrals occurred)
    - nurse level (clinics were carried out and template items were reinforced)
    - patient level (patients received guidebooks, attended nurse-led clinics and improved their physical activity).
- On the basis of these findings, a wider implementation exercise is taking place in collaboration with commissioners, which incorporates changes and refinements to the MOSAICS model to tackle problems of generalisability (e.g. by using briefer, more focused health professional training).

The clinical effectiveness study

This did not show evidence of improved levels of pain and disability. Potential explanations include:

- The evidence base of clinical effectiveness in NICE guidance may not be strong enough to achieve change in clinical end points in practice, or some of the aims of the guidance (e.g. provision of satisfactory information for patients) are worthwhile and necessary in their own right regardless of effect on clinical measures of outcome.
- The implementation was not complete enough to achieve changes in clinical outcomes. There was evidence (quantitative and qualitative) for this in terms of:
    - content of consultations (e.g. exercise advice may have encouraged muscle strengthening at the expense of aerobic exercise such as walking)
    - numbers of patients engaged in full implementation – only a proportion of patients were engaged with the full template or referred on to the nurse and the trial may have lacked power to show an effect at the group level
    - heterogeneity of this patient group
    - resources – the primary care package may be insufficient to achieve patient behavioural change (exercise and weight reduction) without more structured or specialised input to support it (e.g. physiotherapists).
- It may be inappropriate to seek more than minor clinical effect sizes in pain and disability in practice. Maintenance of activity and participation, despite continuing pain, may, for example, be a better long-term measure of effect.

The future

- Evidence from MOSAICS of how to implement the NICE OA guidance in practice is now driving wider dissemination projects (e.g. *Chapter 3, Study 6.2: extending the MOSAICS intervention to new practices via the new NHS*). These provide an opportunity to adapt MOSAICS to generalisable formats and to improve on those areas listed above which limited its implementation.

- Given the findings that patients with an OA diagnosis had better indicator outcomes and that the GP task of giving a diagnosis of OA did not improve after training suggests future implementation and research should address this component of the consultation.
- There is also a need in future research to reflect critically on whether or not:
    - pragmatic trials of clinical effectiveness of NICE guideline implementation are appropriate compared with a combination of implementation studies and new efficacy studies to strengthen the evidence base for the guidelines
    - OA outcomes in practice need to be concerned with aspects over and above improving short-term levels of pain and disability.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Chapter 4 Workstream 3: improving the effectiveness of exercise for osteoarthritis – the BEEP trial

Parts of this chapter have been reproduced from Foster *et al.*[188] © 2014 Foster *et al.*; licensee BioMed Central Ltd. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly credited. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated. Parts of this chapter are reproduced with permission from Moore AJ, Holden MA, Foster NA, Jinks C. Keele University. 2017.

Abstract

Background: Exercise is consistently recommended for older adults with knee pain related to OA, but the effects are typically small and short term. Individualising interventions and supporting adherence are possible ways to improve outcomes of exercise programmes delivered by physiotherapists.

Aim: To test clinical effectiveness and cost-effectiveness of an individually tailored exercise (ITE) programme and a targeted exercise adherence (TEA) intervention compared with usual physiotherapy care (UC).

Methods: A multicentre, pragmatic, parallel group, individually randomised controlled trial, with embedded longitudinal qualitative interviews (ISRCTN93634563). Adults in primary care aged ≥ 45 years with knee pain in one or both knees were recruited through medical records, population survey or referrals to physiotherapy. Participants were randomised to one of three physiotherapy programmes: UC, ITE (individualisation of a lower limb exercise programme) or TEA (support to adhere to longer-term general physical activity). Primary outcomes were pain and function measured by the WOMAC index. Semistructured interviews with a subsample of 30 participants (10 per treatment group) were undertaken soon after treatment ended and then again 12–18 months later.

Results: A total of 526 participants were randomised; 87% provided 6-month data and 79% provided 18-month data. Changes in pain and function were no different between ITE or TEA and UC at 6 months [WOMAC pain UC 6.4 (SD 4.0), ITE 6.4 (SD 4.0), TEA 6.2 (SD 3.8), and WOMAC function UC 21.4 (SD 14.1), ITE 22.3 (SD 13.7) and TEA 21.5 (SD 13.2)] or at other time points. All three groups improved over time, with participants reporting greater control over their knee problem at 6 months than at baseline. UC was less costly than ITE and TEA, yet was as effective and is therefore the dominant treatment. Thematic analysis of interview data highlighted therapeutic alliance between patient and therapist. Physiotherapists' attitudes, beliefs and clinical behaviours about exercise for chronic knee pain changed in line with the specific training package they received, although some reverted by 12–18 months.

Conclusion: Physiotherapy providing individualised interventions or support for exercise adherence did not improve pain and function compared with usual physiotherapy advice and exercise. UC was cost-dominant, meaning it remains the treatment of choice in the NHS.

Trial registration: Current Controlled Trials ISRCTN93634563.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Introduction

Knee pain in older adults is a common disabling problem, managed in the UK mostly in primary care.[189] OA is the most likely underlying diagnosis and refers to a clinical syndrome of joint pain accompanied by varying degrees of function limitation and reduced QoL.[190] The most commonly affected peripheral joint is the knee and OA has been shown by radiography to be present in 70% of community-dwelling adults aged ≥ 50 with knee pain.[190] Structural changes before radiography are common in the remainder.[191] Clinical trials and systematic reviews consistently show the benefit of exercise, in a variety of forms, for this patient group.[33,122,192–194] Exercise improves muscle dysfunction and reduces pain and disability without exacerbating joint damage.[195] It can reduce the risk of other chronic conditions[196] and improve the physical status of people with OA.[197,198] Physiotherapists are the largest group of exercise advisors for musculoskeletal problems in the UK NHS and are therefore an appropriate group with whom to develop and test strategies to improve outcomes from exercise with older adults experiencing knee pain. Previous studies have shown that older adults with knee pain and the physiotherapists involved in their treatment have concerns about the safety of exercise, do not consider exercise as an effective treatment for pain, do not focus on issues of exercise adherence and fail to translate traditional lower limb-focused exercise into sustainable lifestyle changes.[199]

Our previous Treatment Options for Pain in the Knee (TOPIK) trial[34] tested two primary care services for knee pain in older adults, an enhanced pharmacy review and community physiotherapy (based on advice and exercise) and compared these with usual GP-led care and advice. At 3 months, pain reduction was three times greater and functional improvement four times greater in those patients randomised to physiotherapy than to GP-led care and advice, and more people obtained clinically meaningful changes than those in the other two groups.[200,201] Patients received, on average, four physiotherapy treatment sessions. There was evidence that the benefits in pain and function declined in the longer term, suggesting that most patients require some form of monitoring or regular access to physiotherapy or exercise supervision for potential ongoing benefit. The Acupuncture, Physiotherapy and Exercise (APEX) trial,[202] which incorporated a more intensive exercise intervention that was supervised and progressed over six treatment sessions, resulted in greater improvements in pain than the exercise programme in the TOPIK trial.[34,203]

Similarly, other recent trials[30,98] and reviews[204] showed small to moderate, short-term reductions in knee pain and disability that are not sustained in the long term.[192] Exercise is clearly worth doing but we need to find out if, and how, the beneficial effects can be maintained over time. This is crucial to the management of a chronic long-term condition like knee OA. Conversations with patients during telephone follow-ups in the APEX trial helped to explain why the effects of exercise might be suboptimal.[202,203] Participants reported misconceptions about exercise in the presence of knee joint damage and pain, difficulty fitting the exercises into a daily routine, overly complex exercise programmes, insufficient tailoring of the exercise programmes for their individual needs and, in some cases, they simply forgot to do the exercises.

These reviews highlight a lack of information about how to optimise exercise for this patient group.[33,193,194] Most studies are short term, use limited measures of adherence and standardised exercise programmes.[205] Adherence, independent of exercise type, may be an important factor in the success of exercise interventions. There is evidence that better adherence to exercise improves pain relief[192] and disability[205] and that the addition of booster sessions may be helpful in maintaining positive effects on pain and function.[204] A UK consensus[32] identified adherence and tailoring of exercise to individuals as important research topics. One trial showed that greater reduction in knee pain occurred in individuals who reported greater exercise adherence but, unfortunately, 'adherence' was not specifically defined.[145] In the general physical activity literature, a Cochrane review concluded that physical activity programmes that include patient goal-setting, individually tailored and written exercise programmes, some professional guidance and ongoing support, may be the most effective approach.[206]

Recent national and international clinical guidelines support the overall effectiveness of exercise in knee OA, placing it as a key component of core treatment in primary care.[9,15,32] However, there is a lack of evidence around the practical aspects of exercise delivery and maintenance, including how to support individuals to continue to exercise in the longer term. UK guidelines from NICE first published in 2008 recommended that future research should test ways to improve adherence with exercise,[15] and the Chartered Society of Physiotherapy ranked testing ways to help patients incorporate exercise behaviours in their everyday life and increase the long-term effects of exercise as the top musculoskeletal research priority in its national priority setting exercise in 2010.[207]

The BEEP trial[188] is a logical consequence of recent primary care trials, systematic reviews and guidelines for knee pain in older adults, which consistently support exercise-based interventions but highlight the short-term, small- to moderate-sized, benefits.[34,192] The overall aim of the BEEP trial was to test, in older adults with knee pain attributable to OA, whether or not pain and function outcomes can be improved through changing the characteristics of the exercise programme in comparison with UC. The research hypotheses were (1) a physiotherapy-led individualised, supervised and progressed lower limb exercise programme would be superior to UC and (2) a physiotherapy-led intervention targeting exercise adherence in the longer term and supporting the transition from lower limb exercise to general lifestyle physical activity would be superior to UC.

Development

In preparation for the BEEP trial, we conducted two studies. First, a Cochrane systematic review summarised the evidence to date about interventions to improve adherence with exercise in patients with chronic musculoskeletal pain.[208] Second, the Attitudes and Beliefs Concerning knee pain (ABC knee) study investigated the exercise attitudes and behaviours of older adults with knee pain in the community ($n = 611$) and physiotherapists involved in managing older adults with knee pain ($n = 538$), to identify potential barriers to, and facilitators of, exercise for knee pain.[27–29] The findings of our Cochrane review, based on 42 trials with over 8000 patients, suggested using a multifaceted programme combining educational and behavioural strategies to enhance exercise adherence, and incorporating individualisation of the exercise, follow-up and supervision to improve exercise adherence.[208] No one theoretical model underpinning exercise adherence was shown to be superior and the interventions that incorporated motivational strategies showed promise. A subsequent review by Bennell and Hinman[209] stated that the optimal exercise dosage is yet to be determined and an individualised approach to exercise prescription is required based on an assessment of impairments, patient preference, comorbidities and accessibility. They suggested that maximising adherence is a key element dictating success of exercise therapy and that adherence can be enhanced by the use of supervised exercise sessions in the initial exercise period followed by home exercises. Bringing patients back for intermittent consultations with the exercise practitioner, or attendance at 'refresher' sessions, was also suggested to assist long-term adherence and result in improved patient outcomes. Our ABC knee study showed that, while physiotherapists use advice and exercise routinely for older adults with knee pain, they do not consider exercise to be an effective treatment for pain, they have worries about its safety, they provide care over relatively few treatment sessions, thus reducing the capacity to adequately individualise, supervise and progress the exercise programme, and they do not routinely follow up patients to check adherence or to support the translation of lower limb-focused exercise into sustainable lifestyle changes in physical activity. Although patients are aware of risk factors for progression of knee pain and OA (e.g. sedentary lifestyles), they find it hard to make and maintain appropriate lifestyle changes and express the desire for more support from health professionals. Marks[210] describes the many factors that influence exercise adherence in patients with knee OA as either intrinsic (personal factors such as self-efficacy, motivation, age, sex and disease status) or extrinsic (e.g. environmental or social factors or other lifestyle issues). These barriers can vary both over time and between individuals, meaning that one single approach to enhancing exercise adherence might not be as effective as an individually tailored approach. The World Health Organization advocates an 'adherence counseling toolkit' that can be used to systematically assess barriers and facilitators to adherence, and suggests that new interventions to enhance adherence are required.[211]

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

In addition to the above, prior to the funding application to the NIHR, we held an Arthritis Research UK-supported trial network meeting with 24 participants (including physiotherapists, rheumatologists, GPs, biostatisticians, a systematic reviewer and two patient representatives) to further discuss and reach consensus about the trial design and the content and delivery of the interventions. Key decisions made at this networking meeting included agreeing:

- for the design to be a three-arm parallel trial design
- to include 18-month and 3-year follow-ups to provide long-term outcome data
- that the UC arm should be UC rather than usual GP care given the results of previous trials and reviews
- not to include group-based treatments given the known challenges of limited NHS facilities to conduct groups/classes (e.g. attendance in group interventions has been low in previous trials and some patients have strong preferences not to be treated within groups)
- to include objective physical activity data collection on a sample of included BEEP trial patients.

We considered and discussed the potential for group-based treatment but we chose one-to-one physiotherapy sessions for the following reasons:

- We wanted our UC intervention to reflect UC in the UK and our national survey of physiotherapists in the UK[199] showed that only 10% of 538 responding physiotherapists used group-based treatments for knee OA patients, whereas 90% treat patients in a series of one-to-one sessions.
- We wanted to ensure that, if one intervention in the trial was offered in one-to-one sessions, then all interventions should offer care similarly in one-to-one sessions so that different modes of delivery of treatment could not bias the results of the trial.
- A key component of the exercise intervention in the ITE and TEA treatment arms was individualisation of the exercise programme, with patients prescribed a programme that was specific to their individual assessment findings, and this was progressed over time.
- There are considerable challenges in terms of access to facilities in the NHS to hold group-based treatment sessions (particularly in community physiotherapy services where the BEEP trial was based), whereas all services have facilities to offer one-to-one treatment sessions (this might help explain the results of our national survey).
- Although there have been good clinical results from class-based treatment in some previous trials (for which they have been shown to be similarly effective when compared with one-to-one care), attendance at classes or group treatments has been lower than in one-to-one sessions (e.g. Hurley *et al.*[123]) as it is almost impossible to offer sufficient flexibility in days and times of classes to suit all trial participants.
- Previous research, including our own qualitative interviews with patients with knee OA in the community, highlighted that some patients have strong preferences not to be treated in groups or classes.[212]

The following section describes the content and rationale for the three BEEP trial interventions.

Interventions development

Intervention 1: usual physiotherapy care

Usual physiotherapy care consisting of advice and exercise was the most appropriate control group for the BEEP trial given that randomised trials[34] and systematic reviews[33] consistently show that interventions that include exercise are superior to those that do not. We did not use an attention control group as this trial was designed explicitly as a pragmatic trial, building on evidence about the effectiveness of exercise interventions. Clinical practice guidelines recommend exercise for all patients with OA.[15] In current UK clinical practice, patients are seen in relatively few treatment sessions and are provided with brief and relatively standardised exercise programmes that do not take into account differences between individual patients. Our ABC knee study described usual UK physiotherapy practice and formed the basis for the BEEP trial protocol for UC.[199]

The BEEP trial protocol for UC consisted of advice and lower limb exercise previously tested and shown to be more effective in the short-term than GP-led care and advice alone.[34] Exercises were selected from an agreed template of commonly prescribed exercises {printed from the commonly used PhysioTools computer software [PhysioTools©, Product ID RG-PT1ENG, General Exercises Second Edition (English), Tampere, Finland]}, including specific lower limb muscle strengthening (non-weight-bearing and weight-bearing) and range of movement or stretching exercises. Patients were to receive up to four one-to-one treatment sessions with a physiotherapist over a period of 12 weeks, during which advice to continue to exercise was provided but individualisation, progression and supervision of the exercise programme was minimal, limited to up to four sessions as in UC. Other interventions used frequently by physiotherapists such as ice therapy, manual therapy and electrotherapy were permitted as per UC and they were recorded on CRFs, but the emphasis of the intervention was on supporting the patient to self-care and to follow the advice and exercise programme at home. This UC protocol matched usual UK practice in that it:

- focused on lower limb-strengthening exercises
- relied on self-report rather than the use of exercise diaries to monitor adherence and progress
- was delivered over few treatment sessions and, thus, had limited opportunity for good individualisation, supervision or exercise progression
- provided advice about general physical activity only
- did not offer refresher or booster sessions following the end of the episode of care.[199]

Intervention 2: individually tailored exercise

Standardised exercise programmes such as those used in the UC group in the BEEP trial do not take into account differences between individual patients. Consequently, individuals will be working at relatively different intensities, which may be too little for some to get a training effect and may be too difficult for others. To be optimally beneficial, exercise should be progressed so that appropriate physical stress is placed on the individual for further improvements to be obtained.[213] Exercise self-efficacy (confidence to exercise despite the knee pain) has been identified as a predictor of exercise behaviour in many populations.[214,215] Supervision of exercise can enhance patients' exercise self-efficacy and self-regulatory skills and reassure them that they can perform the exercises. It also provides good opportunity to reassure the patient about pain responses to specific exercises and to change the exercise prescription to ensure optimal performance or more tolerable pain response. Lack of adequate individual tailoring, supervision and progression of the lower limb exercise programme may, in part, explain the small benefits seen in some previous exercise trials. Our previous trial results suggested that a lower limb exercise programme that was supervised and progressed over six treatment sessions resulted in greater improvements in pain than one that was provided over an average of four sessions.[34,203] Therefore, the protocol for ITE was developed to ensure the prescription of an individualised, supervised and progressed lower limb exercise programme.

The BEEP trial protocol for ITE consisted of a supervised, individually tailored and progressed exercise programme. The aim of the intervention was to initiate and progress an ITE programme, which was supervised in clinic, practised at home and progressed in terms of intensity over 12 weeks. The intervention was modelled on a previously successful exercise intervention from one of our previous trials[203] and focused on individualised lower limb muscle strengthening (non-weight-bearing and weight-bearing), range of movement or stretching exercise and balance exercise. The patient and physiotherapist were expected to agree a lower limb exercise programme and targets were to be reviewed and progressed. Individualisation was based on the findings of the physiotherapy assessment of each individual, including biomechanical and physiological observations, pain responses to specific exercises, starting levels of strength, range of movement and balance. Exercises were prescribed for each individual and participants were given their own individual print out of their specific exercise prescription (selected and printed from PhysioTools computer software), and these exercise prescriptions (and print-out instructions) were expected to change over time as the exercise programme was progressed. Physiotherapists were to encourage exercise behaviour change using self-monitoring through use of the BEEP trial lower limb exercise diary to record their adherence with their lower limb exercise prescription (see trial protocol[188]). In order to provide

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

greater opportunity for individualisation, supervision and progression of exercise, patients were to receive between six and eight one-to-one treatment sessions with their physiotherapist. Physiotherapists were to provide advice about general physical activities and there were no scheduled follow-ups (refresher or booster sessions) with the physiotherapist beyond 12 weeks.

Intervention 3: targeted exercise adherence

Adherence to long-term treatment regimes, particularly those involving behavioural components such as exercise, is consistently lower than adherence to medication.[15] Exercise adherence, irrespective of exercise type, may be a key factor determining the success of exercise. In a previous trial of exercise for this patient population, we observed high exercise adherence rates in the short term through to the end of treatment at 12 weeks, which fell to just over 50% at 12 months.[203] Long-term adherence to lower limb exercise is perhaps unrealistic, as these can be challenging to incorporate into daily routines and individual lifestyles, and many patients do not consider them enjoyable, stopping them once symptoms reduce or resolve. In order to support engagement in and adherence to exercise, it may be more realistic to target general physical activities that individuals have previously engaged in and enjoyed (as previous exercise behaviour is, theoretically, the strongest source of self-efficacy information),[215] or physical activities that individuals have positive expectations about (as these positive expectations may increase exercise intentions and ultimately exercise behaviour). Our previous ABC knee study[216] highlighted the many different barriers to and facilitators of exercise and physical activity, and that no single exercise type or exercise setting is acceptable to all. Thus, we designed this TEA intervention to include an adherence-enhancing 'toolkit' of optional tools and techniques for physiotherapists to use with different participants, based on their assessment of individual participants and early feedback from participants (see trial protocol[188] for a summary of the contents of the toolkit). The content of the intervention was directly informed by the results of our Cochrane systematic review[208] and a networking meeting supported by Arthritis Research UK, during which national experts and patient representatives agreed the intervention. Although the general physical activity that was identified and encouraged by physiotherapists was individualised for participants, we anticipated that many may choose walking as it is seen as inexpensive and accessible. Pedometers have been shown to increase step counts in older people[217] and to generally increase physical activity[218] and, therefore, we included pedometers within the suite of options for physiotherapists to give participants who wished to target increases in walking activity.

The BEEP trial protocol for TEA began with a focus on the lower limb (as in the ITE group) but transitioned to focus increasingly on general physical activity adherence over time. In addition to prescribing an individualised, progressed and supervised lower limb exercise programme, physiotherapists were trained to assess patients' current general physical activity levels, their intentions to increase their physical activity levels and their attitudes to exercise for knee pain and general health, as well as exploring their individual barriers to and potential facilitators of exercise. This group was to receive four treatments up to week 12 and a further four to six contacts from week 12 through to 6 months (a total of 8–10 treatment contacts). In the first four treatments, the aim was for participants to initiate and progress an exercise programme with supervision that would include lower limb and general exercise, and to identify (with the support of the physiotherapist) general physical activity opportunities within the local community that were suitable for, and of interest to, the individual. Proactive follow-up from the physiotherapist from week 12 through to 6 months was expected, using choices of telephone and face-to-face contact, providing an additional four to six contacts with each participant. The aim was to enhance long-term exercise adherence, promote increased general physical activity and encourage participants to develop an exercise and physical activity 'habit', shifting the focus away from lower limb exercise in the earlier treatment sessions and towards sustainable lifestyle changes in physical activity in later treatment sessions. The follow-up sessions were designed to ascertain and promote exercise adherence and increases in physical activity levels, support patients to integrate exercises into their activities of daily living, allow repetition and amendment of the exercise programme based on individuals' experiences or concerns, and support progression at a pace suitable for the individual. The adherence enhancing 'toolkit' contained different educational, behavioural and cognitive–behavioural tools and techniques for facilitating physical activity behaviour change, selected for use based on an individualised assessment of each patient. As our Cochrane review identified that no one model of behaviour change was

superior for facilitating adherence to exercise for chronic musculoskeletal pain, different theoretical models underpinned the development of the toolkit, including self-efficacy[214] and self-regulation theory.[219] Educational tools included written education material on a range of relevant issues (including healthy eating and pain medication) and individualised PhysioTools exercise sheets. Behavioural strategies included 'How exercise/physical activity feels' and 'How to monitor heart-rate' guides to facilitate physical activity intensities at moderate levels, self-monitoring through use of physical activity diaries,[188] heart-rate diaries, a visual feedback chart, and graded activity sheet, reminder post-cards for physiotherapists to post to patients, and pedometers to support increases in walking. Cognitive–behavioural strategies included guides to questions to elicit patients' health-related beliefs; assessment of patient barriers and motivators; several 'rulers' to use with patients to assess their readiness; perceived importance and confidence to engage in physical activity; decisional balance sheets; specific, measurable, achievable, realistic, time-related (SMART) goal-setting and behavioural contracting for physical activity to facilitate the translation of intention into physical activity behaviour;[212] and example templates to discuss and generate an individual exercise set-back plan. Finally, physiotherapist-generated information about local physical activity opportunities and facilities in the community were developed for each treatment site, with a focus on easy to access and inexpensive options. The target by the end of the 6-month period was that participants would be engaged in physical activity opportunities within their locality and have had support from the physiotherapist to overcome initial problems or barriers in engaging in these activities. Therefore, the emphasis was on maintenance of physical activity beyond the period of support from a health professional and NHS-based programme.

BEEP trial physiotherapy training programme

In the main BEEP trial, a total of 53 physiotherapists from 11 NHS clinics in the West Midlands and Cheshire Physiotherapy Research Network were trained to deliver one of the three interventions within the BEEP trial: UC, ITE and TEA. This training programme was finalised following the delivery of a draft training programme in the pilot study to 11 physiotherapists from six NHS clinics.

Content of the training programme

The content of the BEEP trial training programme was informed by our previous national surveys of physiotherapists and of older adults in the community with and without knee pain, and our Cochrane review (described in *Development*). It primarily focused on the rationale and specific content of each intervention, but also included practical information for physiotherapists about treating patients within RCTs [e.g. the importance of blinding, working to intervention protocols, prevention of contamination and summaries of procedures for serious adverse event (SAE) reporting].

The training programme was 'stepped', in that all physiotherapists attended the first day and received an update about OA, based on the NICE OA guidelines.[15] Key components of the first day included the primary care (clinical) diagnosis of knee OA, the central role of exercise as a 'core' treatment for older adults with knee pain, current physiotherapy practice for knee pain in older adults and OA, and the comparison between current practice and guidelines.[15,32,199,220]

The second and third days were attended by physiotherapists delivering the ITE and TEA interventions and focused on how to improve outcomes from exercise for older adults with knee pain. The importance of individualisation, progression and supervision of lower limb exercise was highlighted from both physiological and psychological perspectives. Exercise self-efficacy was discussed as an important predictor of exercise behaviour[214,215] and emphasis was placed on the importance of 'selling' exercise to patients. Tools to facilitate physiotherapists to individualise, supervise and progress exercise were provided and practised, including developing written individualised exercise programmes using PhysioTools computer software and use of lower limb exercise diaries.

The fourth and fifth (final) days were only attended by physiotherapists delivering the TEA intervention and focused on the importance of exercise adherence, the physiotherapist's role in facilitating exercise and physical activity behaviour change, and shifting from a lower limb exercise programme to physical activity that might be more likely to be sustained in the long term. A number of behavioural models were drawn

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

on within the TEA training programme, including self-efficacy[214] and self-regulation theory.[219] Each physiotherapist was provided with their own 'adherence enhancing toolkit' containing multiple strategies to be used on an individualised basis with patients to facilitate and sustain physical activity behaviour change (described in *Interventions development*). Use of each strategy was practised through role play and participating physiotherapists were encouraged to practise using the tools with patients in their routine clinical practice, prior to the commencement of BEEP participant recruitment.

The training programme included lectures, interactive workshops, role play, group discussion, problem-solving and case studies, with homework set to consolidate learning. In addition, approximately 10 months after the training programme, all physical therapists were invited to attend a half-day workshop to share best practice and discuss any challenges faced with other physical therapists delivering the same intervention. This was designed to help support physiotherapists to deliver the treatments in line with the intervention protocols. A total of 26 physiotherapists attended this additional workshop.

Evaluation of the training programme

We investigated whether or not taking part in the BEEP trial training programme increased physiotherapists' confidence in managing older adults with knee pain and altered their intended clinical behaviour and their attitudes and beliefs about the role of exercise for chronic knee pain, and explored whether or not the more intensive training for physiotherapists delivering the ITE and TEA interventions led to greater changes in these variables than in those physiotherapists delivering UC.

All physiotherapists were asked to complete a questionnaire before (pre training), immediately afterwards (post training) and after delivering the BEEP trial exercise interventions (post intervention, approximately 12–18 months after their training). The content of the questionnaire was based on our previous national survey of physiotherapists. It included a vignette describing a 'typical' patient $\geq$ 45 years with knee pain, associated clinical management questions, attitude statements about the potential benefits of exercise against which physiotherapists rated their level of agreement, and a measure of PSC in diagnosing and managing older adults with knee pain (adapted from a scale developed for back pain[161]). In order to fully describe participating physiotherapists, the pre-training questionnaire also included questions on their clinical characteristics (e.g. number of years in practice and previous training undertaken on OA and exercise therapy).

Analysis

Descriptive statistics were used to describe the sample as a whole, compare results over time (pre training, post training and post intervention) and between the groups of physiotherapists trained to deliver each BEEP trial intervention. Level of agreement with each attitude statement about the role of exercise for knee pain was determined by the percentage of respondents who 'largely' or 'totally agreed'. We defined consensus as follows: 100% = unanimity, 75–99% = consensus, 51–74% = majority view, and 0–50% = no consensus, in line with our previous research.[216,220] Median scores were provided for the measure of physiotherapists' self-confidence in diagnosing and managing knee pain in older adults, as data were positively skewed. Changes over time and differences between groups of physiotherapists (i.e. those trained to deliver each BEEP trial intervention) were not tested for statistical significance owing to the relatively small number of participating physiotherapists overall.

Results

In total, 52 out of the 53 physiotherapists returned the pre-training questionnaire, 44 (83%) returned the post-training questionnaire and 39 (74%) returned the post-intervention questionnaire. The majority of physiotherapists who took part in the training programme were female (63%), worked exclusively within the NHS (67%), had treated at least one patient $\geq$ 45 years with knee pain per week (62%), and had not received previous post-graduate training in the field of chronic knee pain (73%) or exercise therapy (62%). The mean number of years of clinical experience was 14 (SD 11 years), but ranged from 1 to 42 years (*Table 22*).

TABLE 22 Questionnaire response rate and characteristics of participating physiotherapists

Physiotherapist characteristics	Total ($N = 53$)	UC physiotherapists ($N = 15$)	ITE physiotherapists ($N = 22$)	TEA physiotherapists ($N = 16$)
Questionnaire response				
Pre training, *n* (%)	52 (98)	15 (100)	22 (100)	15 (94)
Post training, *n* (%)	44 (83)	12 (80)	17 (77)	15 (94)
Post intervention, *n* (%)	39 (74)	13 (87)	15 (68)	11 (69)
Number of years in practice				
Mean (SD; range)	13.5 (11.1; 1, 42)	18.1 (12.0; 1, 32)	13.8 (12.1; 2, 42)	8.9 (6.4; 1, 23)
Male, *n* (%)	19 (37)	6 (40)	9 (41)	4 (27)
Work exclusively in NHS, *n* (%)	35 (67)	12 (80)	14 (64)	9 (60)
Number of patients usually seen aged ≥ 45 years with chronic knee pain				
< 1 per month, *n* (%)	4 (8)	1 (7)	1 (5)	2 (13)
≥ 1 per month, *n* (%)	15 (30)	7 (47)	3 (15)	5 (33)
≥ 1 per week, *n* (%)	31 (62)	7 (47)	16 (80)	8 (53)
Received postgraduate training in chronic knee pain				
Yes, *n* (%)	14 (27)	5 (33)	4 (18)	5 (33)
Received postgraduate training in exercise therapy				
Yes, *n* (%)	20 (38)	3 (20)	12 (55)	5 (33)

Individual items may not add to totals due to missing data.

Physiotherapists who did not return their questionnaires post training and post intervention were similar in terms of the characteristics that were assessed to those who did complete and return their follow-up questionnaires. Several had moved jobs or had maternity leave during the lifetime of the BEEP trial.

Self-confidence in managing older adults with knee pain

Before the training programme for the BEEP trial, physiotherapists had a median self-confidence score of 8 on a scale of 4–20 [interquartile range (IQR) 7–9]. Scores reduced post training and post intervention (i.e. self-confidence increased) in physiotherapists who received the ITE and TEA training packages, but remained the same in physiotherapists who received UC training (*Table 23*).

TABLE 23 Physiotherapists' confidence in diagnosing and managing older adults with knee pain[a]

Intervention group	Pre training	Post training	Post intervention
UC	8 (8–10)	8 (6–8)	8 (5–8)
ITE	8 (7–8)	6 (5–7)	5 (4–7)
TEA	8 (5–10)	6 (4–7)	4 (4–6)
Total	8 (7–9)	6 (5–8)	5 (4–8)

a All scores are median (IQR) measured by the PSC scale, developed for use regarding patients with low back pain, adapted for use regarding older adults with knee pain.[122] This is the sum of four items of self-confidence in diagnosing and managing older adults with knee pain, measured on a 5-point Likert scale. Lower scores indicate greater confidence (potential range: 4–20).

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Intended clinical behaviour

Treatment approaches Pre training, the most commonly reported treatment approaches for the vignette patient were exercise therapy (100%), heat/ice (69%), manual therapy (29%), acupuncture (15%) and electrotherapy (12%). Post training, the reported use of exercise and heat/ice remained the same, but the use of manual therapy, acupuncture and electrotherapy all reduced (to 27%, 5%, 2%, respectively). This pattern was maintained post intervention 12–18 months later (to 10%, 5%, 8%, respectively) and was evident across all groups of physiotherapists (data not shown), highlighting that participation in the BEEP training programme appeared to effect a shift to a clearer focus on exercise therapy without the addition of other treatment approaches.

At all time points, nearly all physiotherapists reported including advice within their treatment package. This commonly focused on weight loss, pacing, use of heat/ice at home and analgesic use. Pre training, only 45% of physiotherapists reported that they would provide advice about increasing general physical activity. Although this increased post training (80%), this increase was only maintained post intervention by physiotherapists who had attended the training programme for the ITE and TEA interventions (*Table 24*).

Use of therapeutic exercise At all time points, local strengthening exercises and flexibility/range of movement exercises were the most common types of exercise reported by physiotherapists for use with the patient vignette described in the questionnaire. Pre training, only 17% of physiotherapists reported that they would prescribe an aerobic training programme and 50% reported that they would prescribe functional task training. Although this increased post training, particularly for physiotherapists trained to deliver the TEA intervention, this was not maintained 12–18 months later (*Table 25*).

Pre training, most physiotherapists reported that they would provide written and verbal advice on home exercises during the case vignette's initial treatment session, but only 65% would supervise the exercise

TABLE 24 Types of advice included for the patient case vignette by participating physiotherapists

Advice type	Pre training	Post training	Post intervention
Any	51 (98)	44 (100)	38 (97)
Weight loss	47 (92)	43 (98)	37 (97)
Pacing of activities	46 (90)	39 (89)	30 (79)
Use of heat/ice at home	44 (86)	42 (95)	33 (87)
Analgesia	41 (80)	39 (89)	31 (82)
Use of walking aids	23 (45)	19 (43)	18 (47)
Use of knee support	6 (12)	3 (7)	4 (11)
Increasing activity level			
UC	8 (53)	10 (83)	7 (58)
ITE	8 (38)	14 (82)	12 (80)
TEA	7 (47)	11 (73)	8 (73)
Total	23 (45)	35 (80)	27 (71)

Data are number (%) unless otherwise stated.

TABLE 25 Types of exercise for the patient case vignette reported by physiotherapists

Exercise type	Pre training	Post training	Post intervention
Local strengthening	52 (100)	44 (100)	39 (100)
Flexibility/range of movement	48 (92)	42 (95)	35 (90)
Proprioception/balance			
UC	13 (87)	9 (75)	11 (85)
ITE	11 (50)	14 (82)	13 (87)
TEA	10 (67)	11 (73)	9 (82)
Total	34 (65)	34 (77)	33 (85)
Functional tasks			
UC	8 (53)	10 (83)	10 (77)
ITE	12 (55)	15 (88)	11 (73)
TEA	6 (40)	14 (93)	7 (64)
Total	26 (50)	39 (89)	28 (72)
Aerobic training			
UC	6 (40)	5 (42)	2 (15)
ITE	1 (4)	5 (29)	3 (20)
TEA	2 (13)	6 (40)	1 (9)
Total	9 (17)	16 (36)	6 (15)

Data are number (%) unless otherwise stated.

programme. Supervision during the initial treatment session increased post training (84%) and post intervention (95%), a pattern that was seen across all three groups of physiotherapists (*Table 26*).

Changes in the delivery of exercise in follow-up treatment sessions differed according to the training package received (see *Table 26*). In those delivering UC, approximately 50% reported that they would provide written information on home exercises at all time points. Although the proportions reporting that they would supervise exercise and deliver verbal advice on home exercise increased post training, this was not maintained post intervention. Within the group of physiotherapists trained to deliver the ITE intervention, the proportions that reported they would supervise exercise and provide verbal and written advice on home exercise increased from pre training to post training, and this was maintained post-intervention. These patterns were also evident in physiotherapists trained to deliver TEA. However, within this group, the proportion of physiotherapists who reported they would deliver verbal advice on home exercise post training declined 12–18 months later (after delivery of the BEEP intervention to trial participants).

Nearly all physiotherapists reported that they would check that the patient was completing their exercise programme, mainly through observing technique of exercise and verbal questioning. Pre training, only 6% of physiotherapists reported monitoring adherence using an exercise diary (UC 0%, ITE 10% and TEA 7%). This markedly increased in physiotherapists trained to deliver the ITE and TEA interventions after the training programme (53% and 60%, respectively) and after the period of delivery of the BEEP interventions 12–18 months later (60% and 73%, respectively). It did not change in the group of physiotherapists who were trained to deliver UC (post training 8%, post-intervention 15%).

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 26 Exercise delivery for the patient case vignette

Action chosen by physiotherapist	Pre training		Post training		Post intervention	
	Initial treatment session	Follow-up treatment session(s)	Initial treatment session	Follow-up treatment session(s)	Initial treatment session	Follow-up treatment session(s)
Written information on home exercises						
UC	15 (100)	8 (53)	12 (100)	7 (58)	13 (100)	7 (54)
ITE	22 (100)	14 (64)	16 (94)	15 (88)	15 (100)	12 (80)
TEA	13 (87)	10 (67)	15 (100)	11 (73)	11 (100)	8 (73)
Total	50 (96)	32 (62)	43 (98)	33 (75)	39 (100)	27 (69)
Verbal advice on home exercises						
UC	14 (93)	8 (53)	11 (92)	9 (75)	11 (85)	7 (54)
ITE	19 (86)	15 (68)	15 (88)	14 (82)	13 (87)	13 (87)
TEA	14 (93)	9 (60)	15 (100)	14 (93)	10 (91)	8 (73)
Total	47 (90)	32 (62)	41 (93)	37 (84)	34 (87)	28 (72)
Supervision of exercises						
UC	11 (73)	11 (73)	10 (83)	12 (100)	12 (92)	10 (77)
ITE	15 (68)	19 (86)	15 (88)	16 (94)	15 (100)	15 (100)
TEA	8 (53)	11 (73)	12 (80)	14 (93)	10 (91)	10 (91)
Total	34 (65)	41 (79)	37 (84)	42 (95)	37 (95)	35 (90)
Refer to exercise class/group	6 (12)	22 (42)	2 (5)	11 (25)	6 (16)	14 (37)
Refer to student/assistant/technical instructor	5 (10)	15 (29)	1 (2)	7 (16)	2 (5)	13 (34)

Data are number (%) unless otherwise stated.

Number and patterns of treatment

Pre training, physiotherapists reported that they would provide a mean of 3.25 (SD 0.46) treatment sessions for the vignette patient. This remained stable over time in the physiotherapists trained to deliver UC, but increased slightly in those trained to deliver the ITE and TEA interventions. Pre training, 40% of physiotherapists reported offering the patient some form of follow-up after discharge from care. This decreased post intervention in those trained to deliver UC and ITE, but increased in physiotherapists delivering TEA. This group also changed how they would offer follow-up. Pre training, no physiotherapist reported offering the vignette patient a telephone follow-up appointment. This increased to 45% post training, which was maintained post intervention.

Attitudes and beliefs about the role of exercise for knee pain

Prior to the BEEP trial training programme, consensus was reached in only 2 out of the 13 attitude statements relating to the perceived benefits of exercise for knee pain: 'exercises are effective for patients if a radiograph shows mild knee OA' and 'exercises are effective for patients if a radiograph shows moderate knee OA'.[161] Consensus was reached post training and then maintained post intervention on seven of the attitude statements. Although consensus was reached post training on the effectiveness of exercises for patients if a radiograph showed severe knee OA, this was not maintained post intervention. Consensus on the statements 'physiotherapists should prescribe general exercise for every patient with chronic knee pain', and 'general exercise is safe for everybody to do' was only achieved by the groups of physiotherapists trained to deliver ITE or TEA, not in the group of physiotherapists delivering UC (*Table 27*).

TABLE 27 Physiotherapists' agreement with attitude statements about exercise for older adults with knee pain [*n* (%) who largely or totally agreed]

Attitude statement	Pre training				Post training				Post intervention			
	Total (*n* = 52)	UC (*n* = 15)	ITE (*n* = 22)	TEA (*n* = 15)	Total (*n* = 44)	UC (*n* = 12)	ITE (*n* = 17)	TEA (*n* = 15)	Total (*n* = 39)	UC (*n* = 13)	ITE (*n* = 15)	TEA (*n* = 11)
Physiotherapists should prescribe local strengthening exercises to every patient with chronic knee pain	37 (71)	11 (73)	13 (59)	13 (87)	34 (77)	8 (67)	15 (88)	11 (73)	34 (87)	9 (69)	15 (100)	10 (91)
Physiotherapists should prescribe general exercise, for example walking or swimming, for every patient with chronic knee pain	30 (58)	9 (60)	10 (45)	11 (73)	32 (73)	8 (67)	13 (76)	11 (73)	29 (74)	8 (62)	12 (80)	9 (82)
Knee problems are improved by local strengthening exercises	36 (69)	11 (73)	15 (68)	10 (67)	39 (89)	10 (83)	17 (100)	12 (80)	32 (82)	9 (69)	13 (87)	10 (91)
Knee problems are improved by general exercise, for example walking or swimming	28 (54)	9 (60)	13 (59)	6 (40)	35 (80)	8 (67)	16 (94)	11 (73)	33 (85)	10 (77)	13 (87)	10 (91)
Local strengthening exercises for the knee are safe for everybody to do	29 (56)	9 (60)	10 (45)	10 (67)	36 (82)	9 (75)	13 (76)	14 (93)	34 (87)	9 (69)	15 (100)	10 (91)
General exercise, for example walking or swimming is safe for everybody to do	22 (42)	7 (47)	8 (36)	7 (47)	30 (68)	8 (67)	10 (59)	12 (80)	28 (74)	7 (54)	13 (87)	8 (73)
It is important that people with chronic knee problems increase their overall activity	30 (58)	9 (60)	11 (50)	10 (67)	37 (84)	10 (83)	13 (76)	14 (93)	33 (87)	9 (75)	14 (93)	10 (91)
Exercises are effective for patients if a radiograph shows mild knee OA	44 (86)	10 (67)	20 (91)	14 (100)	39 (89)	9 (75)	16 (94)	14 (93)	36 (92)	12 (92)	15 (100)	9 (82)
Exercises are effective for patients if a radiograph shows moderate knee OA	41 (79)	9 (60)	18 (82)	14 (93)	39 (89)	9 (75)	16 (94)	14 (93)	36 (92)	12 (92)	15 (100)	9 (82)
Exercises are effective for patients if a radiograph shows severe knee OA	21 (40)	4 (27)	10 (45)	7 (47)	34 (77)	6 (50)	15 (88)	13 (87)	27 (69)	6 (46)	15 (100)	6 (55)
Exercise works just as well for everybody, regardless of the amount of pain they have	7 (14)	4 (27)	1 (5)	2 (13)	20 (45)	6 (50)	6 (35)	8 (53)	18 (46)	3 (23)	9 (60)	6 (55)
Increasing the strength of the muscles around the knee stops the knee problem getting worse	11 (21)	6 (40)	3 (14)	2 (13)	16 (36)	5 (42)	6 (35)	5 (33)	18 (46)	6 (46)	8 (53)	4 (36)
Increasing overall activity levels stops the knee problem getting worse	6 (12)	6 (40)	0 (0)	0 (0)	18 (42)	5 (42)	6 (38)	7 (47)	19 (49)	4 (31)	9 (60)	6 (55)

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Discussion

We aimed to determine whether or not taking part in the BEEP trial training programme increased physiotherapists' confidence in managing older adults with knee pain and whether or not it altered their intended clinical behaviours, attitudes and beliefs about the role of exercise for chronic knee pain in line with the specific training package that they received.

Prior to the BEEP training programme, participating physiotherapists were already quite confident in their ability to diagnose and manage older adults with knee pain. However, as seen in our previous national survey of UK-based physiotherapists,[201] there were some disparities between their intended clinical behaviour and current clinical guidelines and exercise recommendations for knee OA.[15,32] For example, although nearly all physiotherapists reported that they would use therapeutic exercise and provide advice as part of their treatment package, exercise was provided over relatively few treatment sessions, local (lower limb) muscle-strengthening training was favoured over general physical activity, there was limited supervision of exercise and there was a lack of focus on exercise adherence. There was also uncertainty about the potential benefits of exercise for knee pain.

Directly after the training programme, there were some key changes in the intended clinical behaviours of all groups of physiotherapists, including those who participated in only 1 day of training to support the delivery of UC in the trial. Even within physiotherapists who participated in the UC training programme, use of additional interventions (such as electrotherapy and manual therapy) alongside the exercise programme reduced, there was a greater focus on general physical activity and functional task training, supervision of exercise increased, and level of agreement with the attitude statements about the role of exercise for knee pain also increased. Therefore, these results demonstrate that taking part in the BEEP trial training programme was associated with changes in physiotherapists' attitudes, beliefs and intended clinical behaviours with regard to exercise for a typical patient case (an older adult with knee pain).

The results showed that changes in attitudes, beliefs and intended behaviours were more pronounced in physiotherapists trained to deliver the ITE and TEA interventions. Within these groups, in addition to the changes described above, the number of treatment sessions that physiotherapists reported that they would provide also increased, more physiotherapists said they would monitor exercise adherence using an exercise diary, and there was a higher level of agreement with more of the attitude statements about the role of exercise for knee pain. In addition, their confidence in managing older adults with knee pain also increased. This may be linked to the more specific, targeted content in the training programme about how to individualise, supervise and progress exercise, and how to support longer-term exercise adherence. It may also relate to the intensity and type of content within the training programme for each group of physiotherapists. For example, the ITE and TEA training programmes included more time (in days) and more interactive sessions than the single-day training programme underpinning UC. Previous research has identified that interactive sessions are more effective than didactic educational techniques at facilitating behaviour change within health-care professionals, including physiotherapists.[221–224] However, even in physiotherapists who had undertaken the training to deliver the ITE and TEA interventions, not all changes in intended clinical behaviour and attitudes and beliefs were maintained over the longer term (12–18 months after completing the training programme), a pattern seen in professional groups within other areas of health care.[225] There may be several reasons for this, including barriers at the level of the patient, the professional and the wider health-care organisation. Multimodal approaches that not only address the exercise attitudes and beliefs and intended behaviours of physiotherapists, but also identify and address other potential barriers to long-term behaviour change in physiotherapists, are likely to be required to achieve greater or more sustained changes in these variables over time.[224,226,227]

Although a large number of physiotherapists participated in the BEEP trial training programme ($n = 53$) and we achieved good follow-up rates on their self-completed postal questionnaires after training (83%) and 12–18 months later (74% response rate), the numbers were too small for robust statistical testing to compare changes in beliefs, attitudes or intended behaviours over time and between groups of physiotherapists delivering the three interventions. In addition, as with all survey research that utilises

patient case vignettes, it is possible that intended practice behaviour in response to case vignettes may be different from actual clinical practice.

Conclusions

Physiotherapists who participated in the BEEP trial training programme to deliver the ITE and TEA interventions reported greater confidence in their ability to diagnose and manage older adults with knee pain. For all groups of physiotherapists, the reported attitudes, beliefs and clinical behaviours about exercise for chronic knee pain changed in line with the specific training package that they received. Some of these changes appeared to be sustained through to long-term follow-up at 12–18 months following the training programme and after the end of the period of treatment for BEEP trial participants, while others reverted back towards pre-training levels. The intensity and content of the training programme delivered seemed to be important, as changes in the survey variables among those physiotherapists delivering the ITE and TEA interventions, who attended longer and more interactive training programmes, were greater and more likely to be sustained. Whether or not these changes in physiotherapists' self-reported behaviour translated into changes in their actual behaviour for BEEP trial participants is determined by analysis of the physiotherapy CRFs in the BEEP trial (see *Interventions delivered and adverse events*).

The BEEP pilot study

In preparation for the main BEEP trial, we conducted a small pilot study in two NHS PCTs to investigate the feasibility of training physiotherapists in the new treatment approaches, the acceptability of the treatments to patients and physiotherapists and to test the processes for the main trial. This section summarises the BEEP pilot study and its objectives, methods and key findings, and highlights the key changes made to the main BEEP trial protocol as a result of the learning from the pilot study.

Objective(s) of the pilot study

The primary objectives of the pilot study were to assess the feasibility and acceptability of the ITE and TEA interventions.

The secondary objectives were to:

- test the training programme for participating physiotherapists and inform decisions about resource allocation of staff to the three arms of the trial
- test ways to assess whether or not the interventions are delivered per protocol
- test ways to assess if the exercise programme is individualised and adherence achieved, using physiotherapy CRFs, attendance at clinic sessions, physical activity questionnaires, exercise diaries and accelerometers
- determine the most appropriate tool for measuring self-reported physical activity in the trial population by comparing the PASE and IPAQ questionnaires
- test the recruitment processes, data collection methods and outcome measures to be used in the trial to ensure they operate as expected and they record sufficient information about the detail and fidelity of the interventions.

Methods

Setting

The pilot study was conducted in two PCTs which included a total of six physiotherapy clinics in the West Midlands and Cheshire Physiotherapy Research Network.

Participants

We aimed to recruit approximately 45 patients in total for allocation to one of the three interventions over a period of 6 months.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Inclusion/exclusion criteria

Patients aged ≥ 45 years with knee pain or stiffness in one or both knees referred from general practice or self-referred to physiotherapy were invited to participate. We deliberately chose not to restrict the pilot trial to people with radiographically diagnosed knee OA to reflect current practice in which treatment choices are made on the basis of symptoms rather than radiography findings. Therefore, we included people typical of those seen in primary care and seeking health care.

Exclusions were patients with potentially serious pathology (e.g. inflammatory arthritis), those with previous knee or hip replacement on the affected side(s), those already on a surgical waiting list for total knee replacement, those who had severely restricted mobility or for whom exercise interventions were contraindicated [e.g. those with unstable cardiovascular disorders (severe hypertension, unstable angina, blood pressure > 220/120 mmHg)], and those who had received an exercise programme from a physiotherapist or an intra-articular injection to the knee in the last 3 months. Normal recreational involvement in physical activity was not an exclusion.

Recruitment and treatment allocation

In the pilot study, we used one method of patient identification and recruitment, which was identifying patients who had been referred to physiotherapy by their GP for knee pain, as this had been successful in one of our previous randomised trials.[2] Potentially eligible patients were screened, by a member of the physiotherapy clinical team, for key study eligibility (aged ≥ 45 years with knee pain or stiffness). If deemed eligible, patients were contacted to ask them if they would like to receive information about the study. Those who received the study information were then telephoned by a NIHR Clinical Research Network (CLRN) nurse to screen further for eligibility. If eligible and willing to participate, patients were asked to sign and date a consent form and return it, along with a baseline questionnaire, to the CLRN nurse. Once written consent was received, the study co-ordinator allocated the patient to one of three intervention arms, following the allocation process for the pilot study developed by the study statistician. Pilot testing of the randomisation procedure was not needed as we have previously tested these procedures in similar trials.[34,203]

The BEEP pilot study interventions

Full details of the trial interventions are provided in *Interventions*. *Table 28* provides a summary of the interventions.

Pilot study training programme for physiotherapists

A comprehensive training programme for participating physiotherapists delivering the three BEEP trial interventions was developed and delivered to 11 physiotherapists between April and May 2009 (full details of the BEEP training programme for physiotherapists are provided in *BEEP trial physiotherapy training programme*). The physiotherapists were asked to evaluate how helpful the training had been. They rated each training session on a scale from 1 to 5 with 5 representing the highest level of satisfaction. All sessions were rated highly (> 4.4). Minor changes only, such as the order and timing of the sessions, were made to the main trial physiotherapy training programme as a result of the pilot study. Full evaluation of the BEEP training programme for the main trial can be found in *BEEP trial physiotherapy training programme*.

Data collection

The acceptability, fidelity and feasibility of the interventions, processes, outcome measures and other data recording methods were assessed using a combination of quantitative and qualitative data. The CRFs were developed and used to capture relevant detail about the fidelity of the interventions, adherence to intervention protocols (for UC, ITE and TEA interventions), the individualisation and progression of the exercise programmes (for the ITE and TEA interventions) and the strategies used to encourage exercise adherence and increased physical activity (for the TEA intervention). Acceptability of the interventions was assessed through semistructured, one-to-one, qualitative interviews with seven patients and six physiotherapists. All interviews were recorded with permission, anonymised and fully transcribed.

TABLE 28 Summary of the BEEP trial interventions

UC	ITE	TEA
Up to 4 sessions	6–8 sessions	8–10 sessions
Over the course of 12 weeks	Over the course of 12 weeks	Over the course of 6 months
Advice and information booklet	Advice and information booklet	Advice and information booklet
Exercises selected from an agreed written template	Individually tailored and written exercise programme, aided by written exercise templates	Written agreed personal plan that includes general physical activity opportunities in local community
Focus on lower limb (local) exercise	Focus on lower limb (local) exercise	Focus on both lower limb and general exercise
Supervised and progressed as per UC	Attention to careful individualisation, progression and supervision	Attention to careful individualisation, progression and supervision
No follow-up after 12 weeks	No follow-up after 12 weeks	Monthly follow-up and monitoring sessions using educational, behavioural, and cognitive–behavioural tools and techniques within an adherence-enhancing toolkit

Reproduced from Foster *et al.*[188] © 2014 Foster *et al.*; licensee BioMed Central Ltd. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly credited. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated.

The majority of interviews were conducted by telephone and the questions sought the perspectives of interviewees about their experiences of being involved in the study and the interventions themselves, the associated paperwork and communications with the study team.

Outcome measures

The primary objectives of the pilot study were to test feasibility and acceptability of the interventions. However, we also collected patient self-report data at baseline (after patient consent) and follow-up at 3 and 6 months later, using self-complete questionnaires. The primary clinical outcome measure was WOMAC.[228] Full information on all the secondary outcomes used in the BEEP trial is given in *Outcome measures*. In the pilot study, two measures of physical activity were included (IPAQ and PASE) in order to determine which single measure of physical activity to use in the main trial, as neither had previously been used in our trials with older adults with knee pain.

Ethics

Approval for the pilot study was obtained from Birmingham East, North and Solihull REC (09/H1206/77).

Analysis

The analysis was exploratory and provided data on the feasibility and acceptability of the interventions and the required changes in the recruitment processes, data collection processes and outcome measures to be used in the main BEEP trial. Interview data were analysed thematically.

Findings

Recruitment and follow-up rates

In total, 53 patients were identified as potentially eligible (31 from South Staffordshire PCT and 22 from Shropshire County PCT) during the pilot study. Of these, 28 were screened as fully eligible, gave written informed consent to participate and were allocated to one out of the three interventions. Thus, the same recruitment method as a previous trial[170] led to only 62% of target recruitment in the pilot study (28/45). The overall consent rate of those potentially eligible was good, at 53% (28/53). The mean age of

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

participants in the pilot study was 61 years (SD 7.3 years) and 56% were female. Twelve participants were allocated an accelerometer to wear at baseline and at 3- and 6-month follow-up time points.

Treatment case report form data

The mean number of treatment sessions for each intervention arm was: UC 3.4 sessions (all patients had between one and four sessions in line with the intervention protocol), ITE 6–8 sessions (two patients did not have the protocolised number of sessions; one patient had four sessions and one had five sessions), TEA 8–10 sessions (only two out of eight patients received the protocolised number of treatment sessions).

Only one protocol violation in terms of intervention content was reported during the pilot study (one patient received acupuncture, yet acupuncture was excluded from the BEEP trial intervention protocols). The treatment side effects that were reported on CRFs included muscle/joint soreness, pain, ache and tenderness ($n = 9$, 38%). These are all expected and transient side effects of exercise programmes. The CRFs also showed that all patients were given education and advice, and more than two-thirds had supervised exercises in the clinic and were provided with a home exercise programme. CRFs showed that the exercise and physical activity diaries were used as planned in the ITE and TEA intervention arms and that participants within these trial arms also received an individualised exercise programme (as per protocol).

Linked qualitative interviews

Thirteen semistructured qualitative interviews were completed after the 6-month follow-up time point in the pilot study. The interview sample comprised six physiotherapists (one male and five female; two from each trial arm) and seven patients (two male and five female; two from the UC arm, three from the ITE intervention arm and two from the TEA intervention arm). Recruitment to participation in the interviews was slower than anticipated. However, only one planned interviewee was not interviewed despite numerous attempts. *Table 29* provides a summary of the interview findings from both the patient and physiotherapist interviews.

Outcome measures

Response to the baseline questionnaire was 96% (27/28). The 3- and 6-month follow-up rates were 71% (20/28) and 57% (16/28), respectively. The primary outcomes (WOMAC pain and function) were well completed at all time points with minimal levels of missing data (function score missing for one person at baseline). Missing data rates for secondary outcomes were also low. We had no previous experience of using the PASE or IPAQ so completion rates were particularly important to assess for these measures. Unfortunately, an ambiguity in the PASE questions in the pilot study made it difficult to distinguish missing data from those reporting that they did no exercise. This made comparison between the IPAQ and the PASE difficult. This ambiguity within the PASE was corrected in the main trial and missing data rates were further monitored in the main trial. The accelerometers provided good levels of data after the data cleaning algorithms were applied, with 83% (10/12) of patients wearing accelerometers on ≥ 5 days out of 7 days, for ≥ 10 hours on each day.

Table 30 provides descriptive information for key outcomes measured at the three time points in the pilot study. Overall, the data showed that knee pain and function reduced from baseline and that a small increase in physical activity occurred (as measured by the PASE). Small changes were also seen in the other outcome measures. Outcomes between intervention arms were not compared in the pilot study.

Key changes from the BEEP pilot study

The key learning from the pilot study was related to challenges in participant recruitment using only one recruitment method (from patients already referred to NHS physiotherapy services with knee pain). As a result, two additional recruitment methods were added to the protocol for the main BEEP trial. The recruitment methods for the main trial are fully described in *Invitation, recruitment, consent and randomisation*.

TABLE 29 Summary of key themes from qualitative interviews in the BEEP pilot study

Theme category	Patient interviews	Physiotherapist interviews
Common themes (all arms)	Perceived benefits from the physiotherapy treatment sessions	Limited time in treatment sessions seen as a limiting factor in embedding exercise in patient routines
	Trust in the physiotherapist	Study instructions were clear
	Believed that they have been given appropriate advice	Study team were approachable and accessible
	Found trial information and instructions comprehensive and clear	Training programme was highly rated
	Pleased to be involved in research	Involvement in research was valued
	Satisfied with contact information and opportunities during treatment but expressed a desire for an information point in case of need after the end of treatment	
Key themes in UC	Did not regard exercise as 'treatment'	BEEP UC treatment reflects normal practice
	Not fully convinced that general exercise (e.g. walking) was beneficial	Increase in confidence from training programme (e.g. about latest research)
	Were now thinking about integrating more activity into their everyday lives	
Key themes in ITE	Accepted exercise as necessary (e.g. walking)	Having more treatment sessions in BEEP trial enabled better assessment of 'what works' for individual patients
	Found ways to integrate exercise into daily routines	Exercise diaries seen as useful aids for patients and physiotherapists as reminders and progress monitors
	Intended to keep up increased exercise levels but not always clear how they would do this	CRFs were felt to be good but would benefit from more space to write comments
	Exercise diaries useful initially but then 'not really needed' because remembered to do the exercises	
	Diaries useful as reminders of progress in early stages when knee painful	
Key themes in TEA	Pedometers felt to be interesting and useful	Greater number of treatment sessions/contacts allowed physiotherapists to be more 'creative' and patient-centred with exercise
	Keen to maintain exercise routines developed during physiotherapy-led treatment sessions	Adherence-enhancing Toolkit contents were rated highly
	Gave examples of behaviour changes already adopted (e.g. cycling 2 miles to provide daily elder care instead of taking car)	Rarely used exercise contracts as felt unnecessary unless patient particularly resistant or with clear motivator
	Certainty that activity is of benefit in maintaining knee mobility and pain reduction	Pedometers perceived as useful
		CRFs were felt to be good but limited room for extended commentary

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 30 Summary of patient outcomes in BEEP pilot study (all intervention arms are combined)

Outcome measure (range of possible scores)	Baseline	3 months	6 months
WOMAC			
Pain (0–20)	7.9 (3.3)	5.3 (3.7)	5.3 (3.9)
Stiffness (0–8)	3.3 (1.8)	2.6 (1.4)	2.4 (1.9)
Function (0–68)	27.1 (9.9)	19.0 (12.9)	16.4 (11.0)
SEE scale (0–10)	5.8 (2.3)	5.8 (2.0)	6.0 (2.0)
OEE			
Positive scale (1–5)	3.9 (0.7)	4.0 (0.5)	3.9 (0.4)
Negative scale (1–5)	2.4 (0.8)	2.1 (0.8)	1.9 (0.6)
PASE (0–400)	181 (115)	263 (126)	247 (125)
IPAQ (mets): median (IQR)	2148 (988–4158)	3959 (1122–8800)	7038 (1998–7998)
QoL (EQ-5D-3L) (–0.59 to 1)	0.6 (0.2)	0.8 (0.2)	0.8 (0.3)
Subsample with accelerometry data (count per minute)	324 (166)	332 (150)	295 (84)

OEE, outcome expectations for exercise; SEE, Self-Efficacy for Exercise.
Figures are means (SDs) unless otherwise stated.
The calculation of the PASE outcome measure was not entirely accurate owing to an error on the pilot questionnaire. This error was corrected for the main trial. WOMAC high score = worse outcome; PASE high score = more active; IPAQ high score = more active; SEE high score = more confident that exercise can be done; OEE positive and negative subscales high score = higher expectations that exercise will be beneficial; EQ-5D-3L high score = better health state; and accelerometry high score = more active.

The treatment CRFs from the pilot study showed that the BEEP interventions were deemed feasible and acceptable to both participants and physiotherapists. UC and the ITE interventions were generally delivered in line with the intervention protocols in terms of the number of sessions delivered and the content of treatment. In contrast, the majority of patients allocated to receive the TEA intervention did not receive the number of treatment contacts initially protocolised (between 8 and 10 contacts, which could be a combination of face-to-face and telephone contacts). This issue was therefore particularly highlighted within the main BEEP physiotherapy training programme, for which the importance of treating patients in line with the trial intervention protocol was emphasised. Overall the training programme in the pilot study was well received and only minor changes were made to the programme for the main trial.

The qualitative interviews in the pilot study established that both patients and physiotherapists were generally positive about the BEEP trial and the interventions. However, it was recognised that there was a need for some minor amendments to the format of the interview paperwork and the procedures for telephone interviews, and more space for writing on the CRFs was also requested by the physiotherapists. These amendments were made for the main trial.

In terms of the outcome measures used within the pilot study, these were generally well completed. However, at 6 months the response rate to the follow-up questionnaire was lower than expected (57%). In response to this we ensured that, for the main trial, minimum data collection procedures were in place at the key timepoints (6, 18 and 36 months), as this has been shown to substantially improve response in our previous trials.[34,203] One of the objectives of the pilot study was to determine the most appropriate outcome measure to assess self-reported physical activity within the main trial. We had no previous experience of using the PASE and the IPAQ, so completion rates were particularly important to assess for these measures. Following the pilot study and consultation with the independent Trial Steering Committee, the PASE was chosen for use in the main trial for three reasons: (1) the scoring is simpler than the IPAQ, (2) it appears easier for respondents to complete (all questions are closed form rather than open) and (3) it correlated slightly better with the objective physical activity data from accelerometers (data not shown).

Aims and objectives

The overall aim of the main BEEP trial was to test, in older adults with knee pain attributable to OA, whether or not pain and function outcomes can be improved through changing the characteristics of the exercise programme in comparison to UC.

The research hypotheses were:

1. a physiotherapy-led, individualised, supervised and progressed lower limb exercise programme would be superior to UC
2. a physiotherapy-led intervention targeting exercise adherence in the longer term and supporting the transition from lower limb exercise to general lifestyle physical activity would be superior to UC.

Secondary objectives were to compare the cost-effectiveness of the two exercise interventions with the cost-effectiveness UC, and to investigate differences in (1) the proportions of participants who could be classified as treatment responders, (2) participants' knee pain-related perceptions and expectations and (3) participants' exercise adherence and physical activity.

A linked qualitative study exploring participants' views and experiences of the interventions was also conducted in order to understand participants' experiences of BEEP treatments and if, and how, the treatments helped address participants' barriers to exercise and physical activity in the short and long term.

Methods

Design

BEEP was a multicentre, pragmatic, three-parallel group, assessor-blind, superiority, individually randomised controlled trial, comparing two physiotherapy-led exercise-based interventions with UC, plus embedded qualitative interviews with a subsample of participants.

Participants were randomised (independently) at each treatment site, using block randomisation at a 1 : 1 : 1 ratio, to UC, ITE and TEA. The aim was to recruit 500 participants over a period of 18 months. Each participant's involvement with the trial was for 36 months, during which time they all had access to usual primary care. The primary outcome was assessed 6 months from randomisation, but no analyses were undertaken until the 18-month follow-up time point.

Setting and participants

Participants were recruited from 65 general practices and related physiotherapy services in the West Midlands and North West regions of the UK. Treatments were delivered within physiotherapy centres in five NHS PCTs within the geographical regions of the PCRN of the West Midlands North and the North West. The practices and treatment centres included a mix of urban/suburban/semirural/rural settings.

Participants were eligible for inclusion if they were ≥ 45 years, had current knee pain and/or stiffness in one or both knees, were able to read and write in English, were willing to participate, able to give full informed written consent and had access to a telephone (for minimum data collection). We deliberately chose not to restrict the BEEP trial to people with radiographically diagnosed knee OA in order to reflect current clinical practice and national guidance,[15] in which treatment choices are made on the basis of symptoms rather than on radiographical findings. Therefore, we included people typical of those seen in primary care. Patients who were recruited through the population survey and the record reviews at participating general practices had to have a Chronic Pain Grade[229] severity of between 2 and 4, determined through a brief postal screening survey (see *Invitation, recruitment, consent and randomisation*). This ensured that participants had a mean level of pain and functional difficulty similar to those patients referred to physiotherapy (determined from our previous trials[34,203]) and that the interventions would therefore be suitable for such patients. Chronic Pain

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Grade 2 to 4 categories have been reported as a clinically significant group of knee pain patients in other studies.[50,229,230]

Exclusion criteria were those with potentially serious pathology (such as inflammatory arthritis, malignancy), those who had a total hip or knee replacement on the affected side, those who were on a waiting list for a total knee or hip replacement, those for whom their knee problem was caused by a recent trauma (sports injury, fall or accident), those for whom exercise interventions were contraindicated (such as those with unstable cardiovascular disorders, severe hypertension, unstable angina or congestive heart failure), those who had received an exercise programme from a physiotherapist or a knee joint injection in the last 3 months, those resident in nursing home accommodation, those so severely physically restricted that they could not get to the physiotherapy treatment centres, and those who had a close family member already participating in the BEEP trial. Normal recreational involvement in physical activity was not an exclusion criterion.

Invitation, recruitment, consent and randomisation

Identification of potentially eligible participants

The key learning point from the pilot study was the need to increase recruitment to identify all potentially eligible participants. Therefore, in the main BEEP trial, participants were identified in one of three ways: (1) from general practice computer record reviews to identify those who have consulted for knee pain in the last 12 months, (2) from a population survey of older adults registered with participating practices and (3) from patients referred from their general practice to physiotherapy services for knee pain. The above three recruitment methods proceeded in parallel until the sample size required was reached, although methods 1 and 2 were operationalised in different GP practices. The recruitment methods are described in more detail in numbers (1), (2) and (3) below. Duplication checks ensured that eligible participants were not invited to the BEEP trial more than once. *Figure 8* shows a summary flow chart of BEEP trial recruitment and reasons for ineligibility.

1. General practice record review
   Members of the PCRN informatics team, contracted to work in the participating general practices, screened computer records for adults aged ≥ 45 years who had consulted with knee pain in the last 12 months at the practice. The electronic screen identified patients based on the 14 most frequent knee pain-related Read codes informed by our previous research.[116] Read codes are the standard clinical terminology system that GPs enter onto their computer systems in the UK, identifying patients' clinical symptoms and diagnosis. In addition, the electronic screening protocol excluded those with potentially serious pathology (e.g. inflammatory arthritis such as rheumatoid arthritis and malignancy) and those in nursing home accommodation. GPs were invited to screen the sample list and exclude those patients whom they considered inappropriate to be invited to participate in the trial. PCRN staff then administered the mailing of a short screening questionnaire to check eligibility and determine the patient's Chronic Pain Grade classification.[201] Patients recruited through this method had a Chronic Pain Grade severity of between 2 and 4. The final section of the brief screening questionnaire asked patients if they would be happy to give their consent for further contact. Patients returned their screening questionnaires to the Arthritis Research UK Primary Care Centre at Keele University and only those who appeared to be eligible and consented to further contact formed the sample that was sent information about the BEEP trial.
2. Population survey
   Older adults (aged ≥ 45 years) registered with participating general practices were posted a short screening questionnaire to identify potentially eligible trial participants. GPs were invited to screen the mailing list and exclude those patients whom they considered inappropriate to be invited to participate in the trial. This screen was used to identify those in the community with knee pain of sufficient severity to be included in the BEEP trial[229] and to exclude those who did not meet the eligibility criteria. Patients recruited through this method had a Chronic Pain Grade severity of between 2 and 4. The final section of the brief screening questionnaire asked older adults registered with the practice if they would be

happy to give their consent for further contact. Those who were eligible and consented to further contact formed the sample that was sent information about the BEEP trial.

3. Physiotherapy service referrals
   Older adults (aged ≥ 45 years) with knee pain and/or stiffness referred by their GP or who self-referred to participating physiotherapy services were first screened by a member of the physiotherapy service team for key eligibility criteria. Those potentially eligible were contacted by a PCRN research nurse to find out if they were willing to receive further information about the BEEP trial.

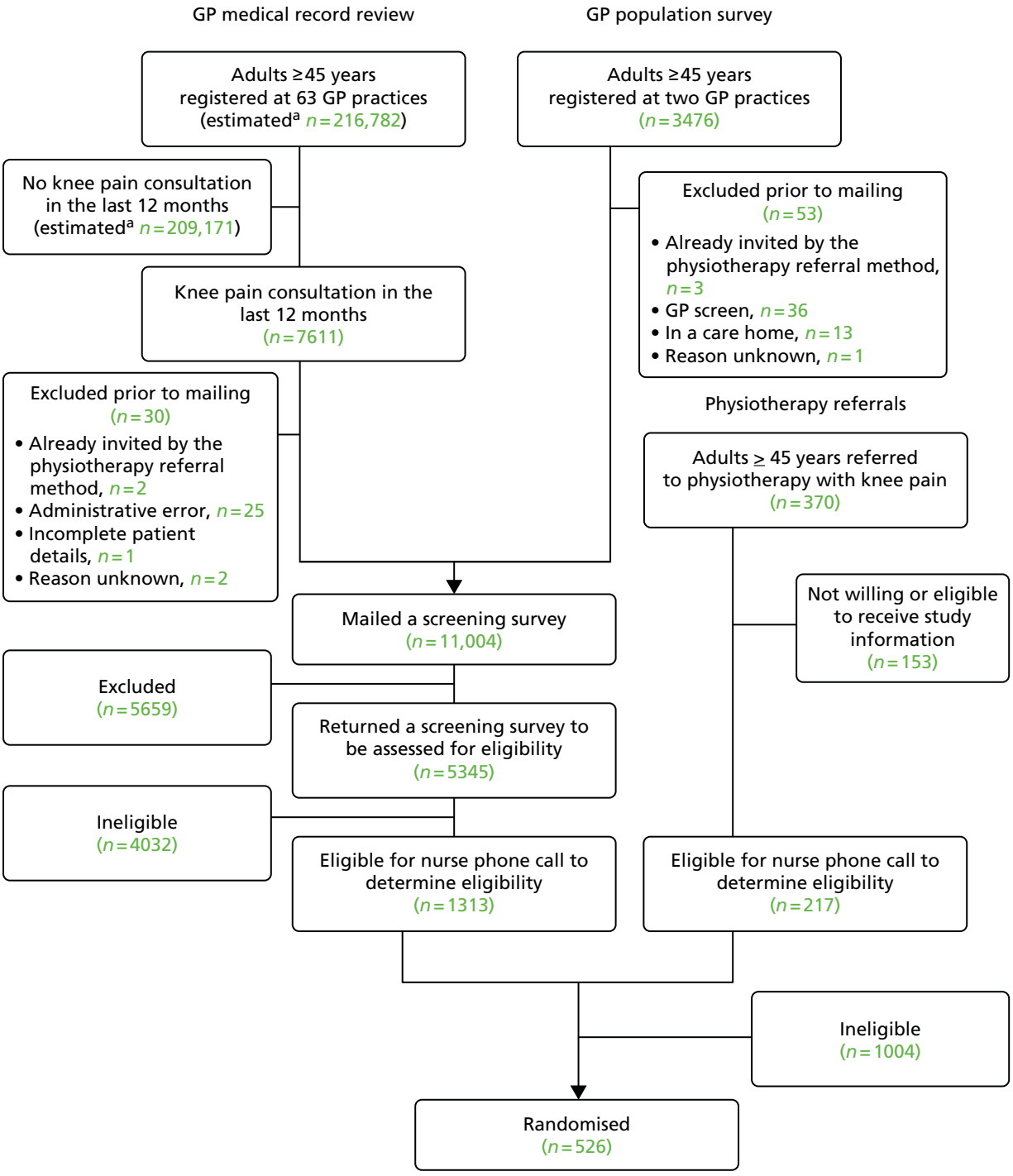

FIGURE 8 BEEP recruitment flow diagram. a, Based on the 2012 practice size data from the NHS Information Centre for Health and Social Care.[231]

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Recruitment and consent of participants to the BEEP trial

From recruitment methods (1), (2) and (3) above, those who met the eligibility criteria and agreed to further contact were posted information about the BEEP trial [a cover letter, a participant information sheet (see trial protocol[188]), a baseline questionnaire, a consent form and a freepost return envelope]. No less than 48 hours after receiving the posted information, a PCRN nurse telephoned the person to further check and confirm eligibility. This check covered the full eligibility criteria in order to ensure that only those who met these criteria were recruited to the trial. In addition, the nurse screened out individuals with known unstable cardiovascular disorders and those who had such severely restricted mobility that they could not get to the physiotherapy clinics for treatment. All potentially eligible participants had the opportunity to discuss the trial with the research nurse prior to deciding whether or not to participate. Those who wished to take part in the trial were asked to sign and date the written consent form, supported by the research nurse over the telephone, and they returned it along with their completed baseline questionnaire to the nurse in a pre-paid envelope. Those who did not wish to take part were asked to indicate this on the consent form and to return it to the nurse in the pre-paid envelope provided. This consent process had been tested in the pilot study and regular audits of the nurse telephone calls formed part of the quality assurance procedures of the BEEP trial. It was possible, based on this consent method, that a very small number of participants in the BEEP trial would be found to be subsequently ineligible, as it was only after consent that participants had a detailed physical assessment by a BEEP trial physiotherapist. Examples of this were expected to include a very small number of participants who had radiating leg pain from a spinal problem or a hip joint problem with referred pain in the area of the knee.

Participating general practices were supported to assist with identification of potentially eligible participants for the BEEP trial through small practice payments to reimburse their time for screening patient lists. Physiotherapy services were supported to participate through financial reimbursement for the time taken out from service delivery for the training programme and additional time for BEEP treatments. Participants did not receive any financial incentives to return screening questionnaires, to take part in the trial or to return their follow-up questionnaires.

Randomisation and allocation concealment

Following receipt of a signed consent form and baseline questionnaire, the BEEP trial administrator randomised the participants using a computer-generated randomisation schedule provided by the musculoskeletal clinical trials unit (CTU) at Keele University, which was password-protected to ensure that research nurses and trial statisticians remained blind to treatment allocation. Participants were individually randomised to one of the three treatment groups in a 1 : 1 : 1 ratio using random permuted blocks of size 3. To ensure that patients at each physiotherapy clinic had a chance of receiving any of the interventions, randomisation was stratified by physiotherapy clinic. Following randomisation, the trial co-ordinator liaised with the appropriate physiotherapy clinic staff to arrange the first appointment for each trial participant. The patient was then informed in writing of the date, time and location of their first appointment in the physiotherapy clinic. The GP of each trial participant was sent a letter to confirm that their patient was taking part in the BEEP trial. Thus, our procedures ensured that baseline data were collected prior to randomisation, that the allocation was concealed until after the patient was recruited into the trial and until the moment of randomisation, and that the person assigning participants to intervention groups (study administrator) had no involvement in the eligibility screen, consent or treatment processes.

Interventions

The interventions were delivered by 47 physiotherapists (15 UC, 17 ITE and 15 TEA) in participating physiotherapy centres in five NHS PCTs. All participants received an advice and information booklet (which included information about the value of specific lower limb and general exercise and simple self-help messages such as the use of analgesics and home heat therapy for pain relief, see trial protocol[188] for a copy of the advice booklet) and a home exercise programme. Different physiotherapists were trained to deliver each of the three interventions. Each of the three intervention groups were detailed previously in *Interventions and development*, with supporting explanation and justification (see *Table 28* for a summary of the interventions). They are all examples of complex interventions as they involved a number of separate,

but interacting, components that were likely to be important to the success of the intervention.[136] Although the key outcomes were knee pain and disability, the interventions were all behaviour change interventions focused on exercise and physical activity behaviour change.

All participants continued to be able to access usual primary care in addition to BEEP treatment. This could include ongoing or new medications, further health-care consultations with other health professionals, and referrals for imaging and surgical opinion, as well as surgical procedures, and these cointerventions were recorded on participants' follow-up questionnaires. For the purposes of the trial, physiotherapy services could contact participants who failed to attend their treatment sessions up to three times in order to try to (re)engage the participant in BEEP treatment sessions. Hydrotherapy, group-based sessions, acupuncture and intra-articular injections were not permitted in any of the BEEP trial treatment protocols. Intervention fidelity was assessed through audits of treatment data collected in trial-specific CRFs (comparing these data with the intervention protocol and with the physiotherapy clinical records). We collected data on the number and content of BEEP treatment sessions using physiotherapy CRFs in all three treatment groups and used those data in our interpretation of the trial results.

Outcome measures

Primary outcomes

This trial had two primary outcomes, lower limb pain and function, measured using the WOMAC Index[228] collected at baseline and all follow-up time points (3, 6, 9, 18 and 36 months). The primary time point was 6 months after randomisation. The primary comparisons were between each of the two new physiotherapist-led exercise interventions (ITE and TEA) versus UC. The psychometric properties of the WOMAC[228] have been extensively studied in knee pain populations in clinical trials of different interventions including exercise[89,232] and the WOMAC has been recently shown to be the most responsive of five pain measures.[233] The pain subscale ranges from 0 (no pain) to 20 (maximum pain) and the function subscale ranges from 0 (no disability) to 68 (maximum disability). It is particularly suitable for the BEEP trial as it specifically captures self-reported pain during activities and the degree of difficulty with everyday physical activities, both of which are key treatment targets of physiotherapy-led exercise.

Secondary outcomes

A range of secondary outcomes were collected: the proportion of treatment responders using the internationally agreed OMERACT clinical responder criteria[200,201] that combine data on pain and function from the WOMAC[228] with patient's global assessment of change (recorded using a 6-point Likert scale); physical activity levels (PASE, which assesses physical activity levels over a 1-week period combining physical activity from several domains including household, occupational and leisure);[169] self-reported BMI (calculated from self-reported height and weight); exercise adherence (attendance at treatment sessions, self-reported adherence to prescribed exercise programme); use of local physical activity facilities in the previous 7 days (single item); a modified version of a measure of treatment acceptability and credibility;[234,235] a measure of illness perceptions (Brief Illness Perceptions Questionnaire[236]); confidence in ability to exercise [Self-Efficacy for Exercise (SEE) scale[237]]; outcome expectations from exercise [Outcome Expectations for Exercise (OEE) scale 2[238]]; anxiety [seven-item Generalised Anxiety Disorder Assessment (GAD-7)[239]]; depression [Patient Health Questionnaire (PHQ) depression scale[240]]; self-reported health-care resource use (both NHS and private health care); and overall health status (EQ-5D-3L[241]). Resource use and EQ-5D-3L data were used in the cost–utility analysis (further details are provided in *Health economic analysis*).

Accelerometry outcomes

Physical activity was also measured in a subsample of participants ($n = 89$ of the 90 initially planned) through snapshots of 7-day accelerometry at each follow-up time point. Accelerometers are motion sensors worn on the hip and were used to estimate physical activity as counts per minute, time spent in light, moderate and vigorous physical activity and proportions of participants who met guideline levels of physical activity.[242] Accelerometers were allocated at the point of randomisation. The accelerometer units were posted out to the participants with full instructions. Participants were asked to wear the unit during

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

waking hours for 7 consecutive days and then to post the unit back to the research centre where the data collected were downloaded and analysed. During the trial, accelerometers were allocated at regular intervals (approximately monthly) to the next three participants randomised to each treatment group. A random allocation procedure was not used as not all participants were willing to wear them. A regular accelerometer allocation procedure was needed to enable accelerometers to be available to collect data (at baseline and follow-up) from a pool of 30 accelerometers available for the trial. The allocation procedure was phased in at the start of the trial to ensure that the system was running smoothly and that accelerometers were being returned in time to be reused by other participants. Those who did not return their accelerometers were posted a written reminder.

Follow-up

Outcome measures were collected by self-report, postal questionnaire before randomisation and at 3, 6, 9 and 18 months (*Table 31*). They are also being collected at 36 months follow-up for further analysis in the future. Therefore, adherence to exercise was measured twice throughout the early phase (0–6 months) and three times during the maintenance phase of the exercise interventions (≥ 9 months). Non-responders were followed up using our standardised CTU follow-up procedures for trials. This comprised a questionnaire, a reminder postcard at 2 weeks and a further copy of the questionnaire at 4 weeks. For participants who did not respond to any of these reminders, we attempted to collect minimum outcome measure data via telephone (by research nurses who were blind to treatment allocation) and by post at key outcome time points (6, 18 and 36 months) in order to try to capture primary outcome data and to minimise missing data. Quality assurance processes within the CTU ensured that training and auditing of the research nurses conducting minimum data telephone calls. In order to reduce participant burden, the 3 and 9 month questionnaires were slightly shorter than those at 6, 18 and 36 months. As direct measures of physical activity may themselves increase physical activity, similar measures of physical activity were used in each intervention group. *Table 31* includes a list of all measures and the time points at which they were collected.

Adverse events

The occurrence of adverse events from all interventions was monitored and assessed using CRFs. An expected minor and transient adverse event from unaccustomed exercise and physical activity is temporary, mild muscle soreness. Physiotherapists delivering the interventions were to advise participants about how to manage such symptoms. Each physiotherapy site was asked to report any SAE experienced by a trial participant immediately to the trial chief investigator that may possibly be related to either the interventions or the trial procedures.

Other data

Other variables collected in participant questionnaires were age, sex, marital status, comorbidities, pain location using a pain manikin, duration of the knee problem, experience of exercise and work status. The following process data were also recorded from the physiotherapy CRFs: number of treatment sessions attended, the main content of each treatment, physiotherapists contact time with participants, the number of treatment withdrawals and treatment non-attendances.

Sources of bias

Selection bias at recruitment was avoided by separating the processes of determining patient eligibility and treatment allocation and by using random permuted blocks overseen by the CTU, thereby not permitting physiotherapists assessing and treating patients to predict the next allocation in their clinic. Trial participants knew they were having physiotherapy-led exercise and had been given brief details about the three interventions in the patient information (see trial protocol[188]). It was not possible to blind physiotherapists but they delivered treatment to participants in only one of the three intervention groups. A PCRN nurse blind to treatment allocation obtained informed consent and oversaw the collection of baseline and follow-up questionnaire data and collected minimum data over the telephone. An evaluation of the success of nurse-blinding procedures was completed and a procedure for reporting incidents in which blinding has been compromised was in place.

TABLE 31 Summary of outcome measures

Data collection	Measurement scale	Time points (months)
Participant characteristics		
Age	Years	0, 3, 6, 9, 18, 36
Sex	Female/male	0, 3, 6, 9, 18, 36
Weight	Stones and lbs, or kg	0, 3, 6, 9, 18, 36
Height	Feet and inches or centimetres	0
Marital status	Married/separated/divorced/widowed/cohabiting/single	0
Work factors		
Current/most recent job title	Free text	0, 6
Currently in a paid Job	Yes/no	0, 6
Working hours	Working full time (≥ 30 hours per week)	6
	Working part time (≤ 29 hours per week)	
Time off because of knee pain including time off to visit any health-care professional	Yes/no (during last 6 months)	6
How many days, weeks or months were you absent from work due to knee problem	Number of days/weeks/months (during last 6 months)	6
Knee problem		
Knee with more discomfort	Tick boxes for left or right	0
Duration of knee problem	In the last 12 months/≥ 1 year but < 5 years ago/≥ 5 years but < 10 years ago/I have had this knee problem for ≥ 10 years	0
Global assessment of change in knee problem	Completely recovered/much better/better/no change/worse/much worse (since the last questionnaire)	3, 6, 9, 18, 36
Knee pain	(WOMAC[228]) 0–20	0, 3, 6, 9, 18, 36
Stiffness	(WOMAC[228]) 0–8	
Function[a]	(WOMAC[228]) 0.68	
Illness perceptions[a]	(IPQ-brief modified for knee pain[236])	
Consequences	0–10	0, 3, 6, 9
Timeline	0–10	
Personal control	0–10	
Treatment control	0–10	
Identity	0–10	
Concern	0–10	
Understanding	0–10	
Emotional response	0–10	

continued

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 31 Summary of outcome measures (*continued*)

Data collection	Measurement scale	Time points (months)
Physical activity and exercise		
Experience of exercise	Personal experiences of exercise	0
Use of local facilities[a]	Use of local facilities for physical activity in the last 7 days	0, 3, 6, 9, 18, 36
SEE	0–10 (SEE scale[237])	0, 3, 6
OEE	1–5 (OEE-2[238])	0, 3, 6
Physical activity[a]	0–400 + (PASE[170])	0, 3, 6, 9, 18, 36
Exercise adherence and treatment credibility	Confidence in, and adherence to, treatment plan	3, 6, 9, 18, 36
Accelerometry[a]	Average counts per minute	For a subsample of participants 0, 3, 6, 9, 18, 36
	Meeting physical activity guidelines (DH Start Active Stay Active[242])	
General health and well-being		
Depression[a]	0–24 (PHQ-8[240])	0, 3, 6, 9, 36
Anxiety[a]	0–21 (GAD-7[239])	0, 3, 6, 9, 36
QoL[a]	–0.59 to 1 (EQ-5D-3L[241])	0, 3, 6, 9, 18, 36
Body manikin (pain)	Body area shaded to represent pain in the previous 4 weeks that had lasted a day or longer	0
Managing your knee problem		
Current medication for knee problem	Prescribed and over-the-counter medications (with dosage)	0, 3
Prescribed medication for knee problem	Prescribed medications in the last 6 months/ 18 months (with dosage and length of supply)	6, 18, 36
Cost of over-the-counter treatments/ appliances	For example, painkillers, anti-inflammatory drugs, TENS machine, hot and cold packs, knee supports (£ in the last 6 months/18 months)	6, 18, 36
Health-care utilisation	Contact with NHS and private health-care professionals, number of visits/inpatient stays, types of investigations/treatments in the last 6 months/18 months	6, 18, 36

IPQ-brief, Brief Illness Perception Questionnaire; OEE-2, Outcome Expectations for Exercise scale 2; PHQ-8, Patient Health Questionnaire 8; TENS, transcutaneous electric nerve stimulation.

a Measure used as an outcome to test for clinical effectiveness (along with BMI and the OARSI responder criteria,[201] not indicated on this table as derived by combining individual measures of height and weight for BMI and WOMAC pain and function with global assessment of change for the OARSI responder criteria).

Reproduced from Foster *et al.*[188] © 2014 Foster *et al.*; licensee BioMed Central Ltd. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly credited. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated.

Data entry, coding, security, storage and management followed the standard operating procedures in the musculoskeletal CTU at Keele University. Data enterers received training in line with CTU procedures and were blind to the identification of the three intervention groups. Random 10% data entry accuracy checks were conducted at regular intervals and data accuracy was audited and accuracy rates recorded. The trial statistician remained blind until after the creation of a locked analysis data set at 18-month follow-up and the completion of the primary and secondary analyses.

We compared available variables between consenting and non-consenting individuals, trial participant withdrawals and completers to evaluate external validity. All participants were free to withdraw from the trial at any time without having to give any explanation. When possible, we collected information about the reasons for withdrawal from the trial. Using validated outcome measures helped to reduce measurement error. Treatment was recorded by physiotherapists in standardised formats and audits of these CRFs were undertaken throughout. Feedback to physiotherapists delivering the interventions was provided, when necessary, so they can try to modify and improve the delivery of the interventions. Each intervention was supported by a specific protocol and documentation, developed for the BEEP trial and previously tested in the pilot study.

Sample size and power

The sample size for the main trial was based on the primary outcome measures: WOMAC pain and function subscales[228] compared at 6 months post randomisation between the UC group and either of the two other intervention groups. The timeframe selected was 6 months because our TOPIK trial[34] showed the benefits from exercise had reduced by this point with UC in comparison with usual GP-led care and advice only. The BEEP trial was powered to detect an effect size of 0.35 for both WOMAC pain and function, an effect size classified as 'small' to 'moderate' using standard benchmarks by Cohen.[81] We chose to power the trial based on a specified effect size (rather than minimum important change) as Terwee *et al.*[243] reported a lack of consensus as to minimum important change on the WOMAC. From our previous trials,[203] we estimated that the SD for WOMAC pain and WOMAC function at 6 months' follow-up would be 5 and 17, respectively. Therefore, an effect size of 0.35 equated to a 1.75-point difference on WOMAC pain and a 5.95-point difference on WOMAC function. To achieve 80% power and a 5% significance level (two tailed), we required 129 patients per treatment group, giving a total sample of 387 patients.[244] Allowing for a 20% loss to follow-up rate (informed by our previous knee pain trials[34,203]), we aimed to randomise a total of 500 participants to the trial over a period of 18 months.

For practical reasons, the sample size was not inflated to allow for clustering of individual patients being treated by the same physiotherapist,[245,246] but rather the trial provides useful estimates of clustering effects and we adjusted for therapists in a sensitivity analysis. We anticipated that a minimum of 36 physiotherapists (12 per treatment group) would be trained to deliver the trial treatments and that each physiotherapist would treat approximately 18 patients. Thus, at the end of the trial, we anticipated being able to estimate an ICC to inform future similar trials.

A formal power calculation was not completed to determine the number of patients to wear an accelerometer. The aim to include 30 per group (a total of 90) was based on practical considerations (restricted by the number of accelerometers available for data collection) and on having a reasonable sample size to analyse the data using parametric statistics.

Statistical analysis

The trial statistical analysis plan was conducted and reported using Consolidated Standards of Reporting Trials (CONSORT) guidelines[245–248] and agreed by the BEEP Trial Steering Committee and Data Monitoring Committee.

Recruitment and follow-up

The numbers of participants identified and recruited using each recruitment method are reported in a flow chart, along with the numbers of participants returning a questionnaire at each follow-up stage. The baseline characteristics of participants are reported for the three treatment groups (to explore the effectiveness of randomisation) and for those with and without data at each follow-up time point (to explore any selective loss to follow-up).

Primary and secondary trial analysis for clinical outcomes

The primary and secondary trial analyses were conducted blind to intervention group by the statistician and an independent statistician verified the analysis of the primary outcomes. The primary analysis compared

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

each of the ITE and TEA interventions with UC on an intention-to-treat basis for the primary outcome at 6 months post randomisation. Secondary analysis included the analysis of the primary outcome at the secondary end points (3, 9 and 18 months) and analysis of the secondary clinical effectiveness outcomes at all follow-up time points. Estimates of clinical effect were derived using analysis of covariance (ANCOVA) for continuous outcome measures and logistic regression for categorical outcomes and are presented as mean or percentage differences (as appropriate) with 95% CIs after adjustment for covariates defined a priori as:

- baseline for the outcome of interest (baseline adjustment is not relevant for the OARSI responder criteria as it incorporates baseline levels of knee pain and function into the follow-up outcome)
- age
- sex
- duration of the knee problem
- physiotherapy treatment centre (as was used in the randomisation algorithm).

A secondary analysis modelled the longitudinal trajectory of WOMAC pain and function over time (the primary outcome) and the PASE physical activity measure using generalised estimating equations. Model predictors included the a priori covariates listed above (with the baseline for the outcome of interest modelled as a covariate rather than outcome as recommended by Peduzzi *et al.*[249]), time, treatment and a time × treatment interaction. A linear model was fitted to the data (with time as a continuous measure) initially; however, if the trajectory over time was non-linear, quadratic or cubic terms were explored. We also explored whether or not conclusions differed if time was represented in the model as a categorical variable. All model estimates were presented with 95% CIs derived using robust standard errors. The primary and secondary trial analyses were conducted after imputation of missing data. Missing data was imputed using the multiple imputation routines in Stata version 12. All primary and secondary clinical effectiveness outcomes (excluding accelerometry) were included in the imputation model and had their missing data imputed. The intention-to-treat analysis included participants who were treatment protocol violators, but did not include participants who were post-randomisation exclusions. Information on any adverse events is reported.

Sensitivity analysis for clinical outcomes

The following sensitivity analyses were completed and are reported.

- Therapist effects: these were explored by adding a random-effect term (to represent the treating therapist) to the main treatment models. The models with and without the therapist effect were compared at each outcome time point to explore the impact of the treating therapist on the effect estimates obtained.
- Imputation of missing data: a complete case analysis was conducted and the results from this compared with those from imputed data.
- Model covariates: treatment models were run on an unadjusted basis to investigate if the inclusion of model covariates had an impact on the main trial findings.
- A per-protocol analysis: this was completed on the subsample of participants judged to have received treatment in line with the specified treatment protocols. Criteria for determining the per-protocol group assignment were established by the Trial Management Group and approved by the Trial Steering Committee before analysis began.

Exploratory subgroup analysis

A specific subgroup analysis with strong theoretical rationale compared the clinical outcomes of pain and function of those who reported high exercise adherence with those who reported lower exercise adherence. We hypothesised that those who reported higher adherence would have improved pain and function outcomes.

Analysis of adherence data

Descriptive statistics were used to examine exercise adherence and were reported as numbers and percentages, or means and SDs, as appropriate. Exercise adherence was measured by reporting the number of treatment sessions attended in each treatment group and by reporting the following measures at each follow-up time point:

- self-reported exercise adherence (measured by agreement with the statement 'I have been doing my exercises as often as I was advised')
- frequency and duration of physiotherapy exercise completed by the participant
- change (from baseline) in level of physical activity (measured by PASE and, in a subsample of participants, accelerometry)
- self-reported use of local exercise and physical activity facilities.

Analysis of accelerometer data

Accelerometry data from the subsample of BEEP trial participants was analysed using ActiGraph (version 6.6.3; Actigraph© Lifestyle Monitoring System, Pensacola, FL, USA)[250] accelerometer software. Prior to analysis, data were cleaned by excluding time periods containing > 60 minutes of zero count (for which it is assumed the accelerometer was not worn) and by including only those participants who had worn the accelerometer for at least 5 days for ≥ 10 hours. Two sensitivity analyses were conducted by changing the threshold to exclude minutes of zero count at 30 and 90 minutes, respectively, with the later threshold recommended for participants with knee pain.[251] For each participant, we generated the following at each data collection point: number of valid minutes, counts per minute and proportion of time spent at each level of physical activity using the cut-off points from Freedson *et al.*[252] We also calculated the proportion of participants meeting exercise/physical activity recommendations.[242] Descriptive statistics are given for the accelerometer variables at baseline and by treatment arm. Change (between baseline and each follow-up time point) in the average count per minute and in the proportion meeting the exercise recommendations were analysed using ANCOVA and logistic regression, respectively, with analyses adjusted for the covariates defined for the clinical effectiveness analysis and by treatment arm.

Health economic analysis

A cost–utility analysis was undertaken to compare ITE with UC and TEA with UC, using QALYs as the measure of benefit. The base-case analysis for the evaluation adopted a NHS and Personal Social Services (PSS) perspective. A complete case analysis and a broader perspective incorporating NHS/PSS costs, patients' personal expenditure and costs associated with work over an 18-month follow-up period was considered as the sensitivity analysis.

Responses from the EQ-5D-3L generic instrument were used to measure preference-based health-related QoL at baseline, 3, 6, 9 and 18 months. For each patient, EQ-5D-3L index scores from responses to the questionnaire were obtained using the UK value set.[253] Using the area-under-the-curve approach that links each patient's utility scores at different time points,[254] EQ-5D-3L responses were used to generate QALYs for each patient.

Knee-related health-care resource use data were collected by patient self-report questionnaires administered at 6 and 18 months. NHS resource use data included primary care contacts (GPs, practice nurses, community physiotherapy), contacts with other health-care professionals (e.g. hospital consultants, physiotherapists and acupuncturists), hospital-based investigations and procedures [e.g. radiography, magnetic resonance imaging scans, knee-related injections and surgical procedure]), and prescribed medications. Additional information, reported in other trial communication by participants, about any knee surgery undertaken or planned within the 18-month follow-up period was also recorded. Non-NHS (health-care) costs were obtained by asking patients about their use of private health care and purchase of over-the-counter medicines and use of local exercise facilities, treatments or appliances. In order to assess broader economic consequences of the interventions beyond health-care resources, self-reported data on occupation and time off work taken because of their knee pain over the 18-month period was also collected.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Details of the number of trial-related physiotherapy sessions attended by each participant were collected through the CRFs. The costs required to deliver the two new interventions included additional physiotherapy sessions, telephone contacts and pedometers supplied to participants in the TEA intervention. The cost of the interventions included an average 45-minute initial assessment and treatment session, followed by 20-minute face-to-face physiotherapy sessions and 7.5-minute telephone contacts in the TEA intervention. Unit costs were obtained from various sources, including the BNF,[174] *NHS Reference Costs*[175] and *Unit Costs of Health and Social Care*[173] and applied to resource use items. All unit costs used in this analysis are reported in *Table 32*, using a common 2012/13 price year.

All analysis was based on the intention-to-treat principle. In the base-case analysis, multiple imputations were used to impute all missing values for the EQ-5D-3L and non-surgery NHS costs for non-responders to the 6- and 18-month questionnaires. The base-case analysis involved imputing only the non-surgery NHS

TABLE 32 Summary of health-care resource unit costs

Health-care resource	Unit cost (£)
Primary care[a]	
GP consultation per 11.7 minutes	34
Practice nurse consultation per hour	44
Nurse home visit per hour	60
Community physiotherapist per hour	30
Secondary care contacts[b]	
Orthopaedic surgeon: first attendance	128
Orthopaedic surgeon: follow-up	102
Rheumatologist: first attendance	202
Rheumatologist: follow-up	133
Acupuncturist: first attendance	49
Acupuncturist: follow-up	44
Physiotherapist: first attendance	49
Physiotherapist: follow-up	44
Occupational therapist: first attendance	75
Occupational therapist: follow-up	68
Podiatrist: first attendance	74
Podiatrist: follow-up	68
Intervention cost	
First physiotherapist session: 45 minutes	22.50
Follow-up physiotherapist sessions: 20 minutes	10
Telephone physiotherapy consultation: 7.5 minutes	3.60
Pedometer	5
Prescribed medication	Patient specific[c]
Medical investigations/Interventions	Patient specific[b]

a Unit Costs of Health and Social Care 2012.[173]
b NHS Reference Costs 2012/13.[175]
c BNF.[174]

costs, in order to account for the additional knee surgery data retrieved for the 18-month follow-up period that were recorded separately from postal questionnaires and trial communication. A complete case analysis was also undertaken and is presented as a sensitivity analysis. Mean differences are all presented with bootstrapped 95% CIs.

An incremental cost-effectiveness analysis was conducted to determine the difference in costs and outcomes between ITE, TEA and UC. The unit of outcome was the incremental cost per QALY gained (incremental costs divided by QALYs) associated with the ITE and TEA interventions. The estimation of cost-effectiveness within the trial was based on the principles of dominance. If an intervention was found to be less effective and more costly than at least one of its comparators, then it was not included further in the economic analysis. To account for uncertainty, non-parametric bootstrapping was used to derive 5000 paired estimates of mean differential cost and QALY scores. These were then presented graphically on a cost-effectiveness plane.[255,256] The analysis also controlled for any possible between-group imbalance in baseline utility using a regression-based adjustment in order to avoid biased estimates of QALY scores.[257] Although costs were collected over an 18-month period, discounting was not applied to the 18-month questionnaire data, as this contained resource use data from both the last 6 months of year 1 and the first 6 months of year 2 and could not be disaggregated. When productivity loss was reported, costs were assigned using the human capital approach; self-reported days of work absence were multiplied by respondent-specific wage estimates identified from annual earnings data and UK Standard Occupational Classification coding.[258,259] Productivity costs were not imputed and a complete case analysis was undertaken. Statistical analysis was performed using Stata v12.

Linked qualitative interviews

Qualitative methods have an important role to play in evaluating complex interventions[136] and in helping to interpret the findings from RCTs.[260,261] In the BEEP trial, linked, longitudinal, qualitative interviews explored participants' experiences of treatment, their views of the acceptability of their treatment, the impact of the interventions on participants' exercise and general physical activity behaviour, and explanations for change in knee symptoms and exercise behaviour over time. The topic guides (see *Appendix 3*) included questions that explored participants' views of their physiotherapy treatment, how they got on with their exercise programme and what factors they felt helped or hindered them to adhere to the exercises following the end of their treatment contact with the physiotherapist and approximately 12–18 months later. We also asked what they felt the future held for their knee problem. Interviews were semistructured, face to face and lasted approximately 1 hour. Interviews were longitudinal in that they were conducted following physiotherapy treatment completion and also in the longer term (12–18 months later, timed to take place after their 18-month BEEP follow-up questionnaire). The number invited was determined by ongoing data analysis and theme saturation. We anticipated continuing until approximately 30 sets of longitudinal interviews were conducted. Purposive sampling based on data collected in the 3- and 6-month follow-up questionnaires ensured a diverse range of characteristics in terms of demographic details (age and sex), intervention group, severity of knee condition determined by WOMAC pain and function scores (as this could be linked to participants' willingness to exercise) and changes in WOMAC scores after intervention completion (as these could be linked to participants' views about success of the BEEP interventions and longer-term exercise adherence). With participants' written consent, interviews were audio-recorded and transcribed verbatim by professional transcribers. All transcripts were anonymised and data management and analysis facilitated by NVivo (version 9; QSR International, Warrington, UK). Earlier interviews (at the end of physiotherapy treatment completion) were analysed and informed the interview content of the interviews at the longer-term follow-up.

We used an emergent and layered approach to analysis which allowed for both induction and deduction.[262] First, using the principles of constant comparison,[184] we open-coded all transcripts. This enabled patient experiences of the interventions, general views on exercise and barriers and facilitators to exercise to be explored. A researcher coded the transcribed data, with a sample of interviews being independently coded by a further three members of the research team to ensure transparency, and agreed emergent themes at successive stages of the data collection and analysis. Second, we applied a more deductive approach by

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

rereading transcripts and allocating data to predetermined codes of individualisation, supervision and progression (three core characteristics of exercise delivery within the BEEP study). Third, a focused within-case and cross-case longitudinal analysis was performed by asking descriptive and interpretative questions of the data.[263] Descriptive questions investigated what changed in terms of participants' knee condition and use of exercise or general physical activity levels and what were key influences on these changes. This enabled an overall understanding of differences from the first interviews (post intervention) to the second interviews (12–18 months after the end of the BEEP intervention) and potential reasons for those differences. Interpretive questions were then used to investigate, for example, which changes inter-relate and how they relate to existing theories of exercise adherence. Data summary frameworks facilitated the identification of patterns across time.[262,263] Qualitative data analysis was undertaken separately to the quantitative data analysis in the first instance to facilitate an interpretative approach and not constrain the analysis by quantitative variables or findings.

Results

Participant recruitment and flow

Figure 8 illustrates the flow of patients during the trial. Out of 1530 potentially eligible participants, 526 adults aged ≥ 45 years were eligible and agreed to participate. A total of 364 were recruited using general practice record reviews, 45 were recruited from the population survey and 117 were recruited from physiotherapy referrals (*Table 33*). The three recruitment methods led to similar patients being included in the trial. Those recruited from physiotherapy referrals were slightly younger on average (60 years) than those recruited from general practice record reviews and the population survey (63 and 65 years, respectively). Approximately half of those patients recruited through the GP record review and the population survey had a chronic pain grade of 2 (moderate pain and disability over the past 6 months).

In total, 526 were randomised (176 to UC, 178 to ITE and 172 to TEA) between October 2010 and February 2012. The group of patients randomised were the same age as all those who were eligible to receive the nurse telephone call, but included slightly fewer women (see *Table 33*). In total, 12 patients randomised in the BEEP trial were found to be subsequently ineligible at their first physical assessment with BEEP trial physiotherapists (one from UC, two from ITE and nine from TEA intervention arms). These 12 patients were all found not to have knee pain related to OA; two were diagnosed with meniscal tears needing onward medical attention, two had referred pain from a back problem, two had referred pain from a hip problem, two had combined hip/back problems that explained their knee pain, three had neurological conditions that explained their knee pain and one had no knee pain by the time they consulted the BEEP trial physiotherapist. We excluded these patients from follow-up and analyses and, therefore, 514 participants form the data set for the BEEP trial. The data entry of the analysis data set had a very low error rate (1 in 10 data checks gave error rates lower than 0.1% at baseline and at each follow-up time point).

There were no important differences between the intervention arms at baseline (*Table 34*). On average, participants were 63 years old, were overweight with moderate pain and functional disability scores and had low anxiety and depression levels; most had been experiencing knee pain for between 1 and 5 years. Overall, they had low physical activity levels and were not meeting recommended guidelines for physical activity, and only about one-third had used local facilities for physical activity in the last 7 days. On average, they strongly believed that their knee problem would last a long time and were highly concerned about it. Despite these views, on average, they were positive about the ability of treatment to help their knee problem and had generally positive expectations about the benefits of exercise specifically.

Interventions delivered and adverse events

Treatment was delivered by 47 physiotherapists in 10 treatment clinics. Out of all 514 in the BEEP trial, 39 (7.6%) participants received no physiotherapy sessions despite several attempts at contact to offer an appointment [$n = 12/175$ (6.9%), $n = 15/176$ (8.5%) and $n = 12/163$ (7.4%) in the UC, ITE and TEA arms,

TABLE 33 Descriptive characteristics of participants recruited by each recruitment method in the BEEP trial

	Recruitment via GP record review of consulters with knee pain in the last 12 months				Recruitment via population survey				Recruitment via physiotherapy referral	
Key characteristics	Consulters mailed survey (*N* = 7581)	Survey responders (*N* = 4092)	Eligible for nurse telephone call (*N* = 1171)	Randomised (*N* = 364)	Mailed survey (*N* = 3423)	Survey responders (*N* = 1253)	Eligible for nurse telephone call (*N* = 142)	Randomised (*N* = 45)	Eligible for nurse phone call (*N* = 217)	Randomised (*N* = 117)
Age (years)[a]	64 (55–73)	65 (57–74)	64 (56–72)	63 (56–70)	60 (52–70)	64 (55–73)	65 (57–74)	65 (57–74)	59 (52–66)	60 (54–66)
Sex[b]										
Females	4158 (55)	2252 (55)	656 (56)	187 (51)	1752 (51)	670 (54)	77 (54)	24 (53)	113 (52)	55 (47)
CPG[a]										
No pain	[c]	202 (5)	[d]	[d]	[c]	633 (52)	[d]	[d]	–	–
Grade I	[c]	1330 (34)	[d]	[d]	[c]	359 (30)	[d]	[d]	–	–
Grade II	[c]	893 (23)(*37*)[e]	514 (44)	170 (47)	[c]	103 (9)(*47*)[e]	70 (49)	22 (49)	–	–
Grade III	[c]	722 (18)(*30*)[e]	352 (30)	113 (31)	[c]	56 (5)(*25*)[e]	38 (27)	13 (29)	–	–
Grade IV	[c]	823 (21)(*33*)[e]	304 (26)	80 (22)	[c]	62 (5)(*28*)[e]	34 (24)	10 (22)	–	–

–, Data not collected; CPG, Chronic Pain Grade.
a Numbers are medians (IQR).
b Numbers are *n* (%).
c Data missing for non-responders so overall frequency not reported.
d Data missing owing to trial entry selection criteria.
e Figures in italics are CPC percentages calculated for participants with grades II to IV only.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 34 Baseline characteristics of participants ($n = 514$)

Key characteristics	UC ($N = 175$)	ITE ($N = 176$)	TEA ($N = 163$)	Overall ($N = 514$)
Age (years)	62 (9)	63 (11)	64 (9)	63 (10)
Female[a]	87 (50)	91 (52)	84 (52)	262 (51)
Married/cohabiting[a]	135 (77)	134 (77)	126 (79)	395 (78)
Currently in a paid job[a]	81 (47)	67 (39)	66 (41)	214 (42)
BMI (kg/m^2)	29.4 (5.8)	29.4 (5.7)	30.0 (5.5)	29.6 (5.7)
Comorbidities				
Heart disease (including high blood pressure, angina, heart failure, stroke and heart attack)[a]	96 (55)	85 (49)	80 (49)	56 (11)
Lung disease (including asthma and bronchitis)[a]	28 (16)	36 (21)	24 (15)	88 (17)
Diabetes[a]	27 (15)	22 (13)	17 (10)	66 (13)
Depression[a]	40 (23)	47 (27)	27 (17)	114 (22)
Osteoporosis[a]	9 (5)	14 (8)	14 (9)	37 (7)
Pain in at least one body site other than the knee[a]	173 (99)	175 (100)	162 (100)	510 (100)
Widespread pain[a,b]	25 (14)	30 (17)	24 (15)	79 (15)
WOMAC				
Pain (0–20)	8.2 (3.2)	8.5 (3.7)	8.5 (3.5)	8.4 (3.5)
Stiffness (0–8)	3.7 (1.7)	3.7 (1.9)	3.8 (1.6)	3.7 (1.7)
Function (0–68)	27.5 (12.0)	27.8 (12.7)	29.1 (12.0)	28.1 (12.2)
Onset of knee problem[a]				
In the past 12 months	39 (22)	46 (26)	42 (26)	129 (25)
> 1 year but < 5 years	74 (42)	60 (34)	67 (41)	200 (39)
≥ 5 years but < 10 years	33 (19)	37 (21)	26 (16)	93 (18)
≥ 10 years	32 (18)	33 (19)	28 (17)	93 (18)
PASE scale (0–590)	176 (80)	175 (87)	180 (83)	177 (83)
Previous personal experience of exercise on a regular basis[a]	115 (66)	102 (59)	101 (64)	318 (63)
Used local facilities or opportunities that involved any form of physical activity in the last 7 days[a]	63 (36)	56 (32)	47 (29)	164 (32)
SEE scale (0–10)	5.5 (2.3)	5.4 (2.3)	5.3 (2.5)	5.4 (2.3)
OEE scale				
Positive subscale (1–5)	3.9 (0.6)	3.9 (0.6)	4.0 (0.6)	3.9 (0.6)
Negative subscale (1–5)[c]	3.6 (0.8)	3.6 (0.9)	3.4 (0.8)	3.5 (0.9)
GAD-7 (0–21)[d]	2.0 (0.0–5.0)	1.1 (0.0–4.1)	2.0 (0.0–4.0)	2.0 (0.0–5.0)
PHQ-8 scale (0–24)[d]	3.0 (1.0–6.0)	3.0 (1.0–5.7)	2.0 (0.9–5.0)	2.3 (1.0–5.8)

TABLE 34 Baseline characteristics of participants ($n = 514$) (*continued*)

Key characteristics	UC ($N = 175$)	ITE ($N = 176$)	TEA ($N = 163$)	Overall ($N = 514$)
Brief Illness Perceptions Questionnaire (IPQR)				
How much does your knee pain affect your life? (0–10)	5.5 (2.1)	5.5 (2.2)	5.5 (2.4)	5.5 (2.2)
How long do you think your knee pain will continue? (0–10)[d]	10.0 (7.0–10.0)	9.9 (7.0–10.0)	9.2 (7.0–10.0)	10.0 (7.0–10.0)
How much control do you feel you have over your knee pain? (0–10)	4.4 (2.7)	4.0 (2.5)	3.9 (2.8)	4.1 (2.7)
How much do you think treatment can help your knee pain? (0–10)[d]	7.0 (5.0–8.6)	7.0 (5.0–8.0)	7.0 (5.0–9.0)	7.0 (5.0–9.0)
How much do you experience symptoms from your knee pain? (0–10)	6.2 (1.9)	6.2 (2.1)	6.2 (2.3)	6.2 (2.1)
How concerned are you about your knee pain? (0–10)[d]	8.0 (6.0–10.0)	8.0 (6.0–10.0)	8.0 (6.8–10.0)	8.0 (6.0–10.0)
How well do you feel you understand your knee pain? (0–10)	5.7 (2.7)	5.9 (3.0)	5.3 (3.3)	5.6 (3.0)
How much does your knee pain affect you emotionally? (0–10)	4.7 (3.1)	4.9 (3.1)	4.9 (3.0)	4.8 (3.1)
Accelerometer data[e] (subsample of trial participants)				
n	21	21	18	60
Average number of counts per minute (measured by accelerometers)[d]	311 (215–390)	253 (177–302)	232 (163–299)	246 (180–329)
Participants meeting physical activity guidelines[a,f]	4 (19)	1 (5)	0 (0)	5 (8)

IPQR, Brief Illness Perceptions Questionnaire; PHQ-8, Patient Health Questionnaire 8.
a Numbers are *n* (%).
b Defined using the Manchester definition of widespread pain.[264]
c Reverse scored (i.e. a higher score on the negative subscale indicates higher expectations of the benefit of exercise).
d Median (IQR).
e Participants are only included if they have worn the monitor for at least 5 days for ≥ 10 hours. Valid time is calculated assuming that any consecutive runs of zero count lasting for ≥ 60-minutes are counted as non-wear.
f Defined as participants completing 150 minutes each week of moderate-intensity physical activity (accumulated in bouts of ≥ 10 minutes) or 75 minutes of vigorous intensity activity spread across the week (adapted from DH Start Active Stay Active[242]). Bouts calculated using a drop time of 2 minutes (ActiGraph Support[252]). Missing days of data are imputed using the average of the average count for days when data are present.
Values are mean (SD) unless otherwise stated.
All figures are based on data after multiple imputation of missing data has been applied (with the exception of the accelerometer data, comorbidity data and pain at other body sites).
WOMAC higher score = worse outcome; PASE higher score = more active; SEE higher score = more confident that exercise can be done; OEE positive and negative subscales higher score = higher expectations that exercise will be beneficial; GAD-7 higher score = more anxious; PHQ-8 higher score = more depressed; IPQR – affects life, higher score = more affected; IPQR – duration, higher score = lasts a longer time; IPQR – personal control, higher score = more control; IPQR – treatment control, higher score = higher belief treatment can control; IPQR – symptom experience, higher score = more symptoms that are more severe; IPQR – concern, higher score = more concerned; IPQR – understanding, higher score = more understanding; IPQR – emotion, higher score = more emotionally affected.

respectively]. Participants in the UC arm had fewer treatment sessions (median 3, IQR 2–4, range 0–8) than those in the ITE intervention arm (median 6, IQR 4–6.5, range 0–9) and the TEA intervention arm (median 7, IQR 4–8, range 0–11). Of participants in the TEA arm who had at least one treatment session, 94 (62%) had one or more sessions that were by telephone call rather than face to face. The median time from the date of randomisation to the first treatment session was 23 days (IQR 15–32).

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Nearly all UC participants received physiotherapy that included assessment and reassessment (98%), education and advice (98%), a PhysioTools exercise sheet (93%), a home exercise programme (96%) and the physiotherapist supervised their exercise programme in clinic (87%). Muscle-strengthening exercise was used for nearly all patients (94%), range of movement or stretching exercise was used for 74% of patients and balance exercises formed part of the programme for 63% of patients. In total, 156 (89%) of the CRFs were judged to be in line with the treatment protocol for UC.

All ITE participants received physiotherapy that included assessment and reassessment (100%) and a home exercise programme (100%), and nearly all received physiotherapy that included education and advice (99%), an individualised PhysioTools exercise sheet (92%) and the physiotherapist supervised their exercise programme in-clinic (99%). Most patients also had their exercise programme progressed following review (87%) and almost three-quarters were given a lower limb exercise diary (74%). Muscle-strengthening exercise was used for nearly all patients (98%), range of movement or stretching exercise was used for 80% of patients and balance exercises formed part of the programme for 76% of patients. In total, 109 (62%) of the CRFs were judged to be in line with the treatment protocol for the ITE intervention.

Out of the 163 participants in the TEA intervention arm, 94 (62.3%) had physiotherapy contact over the telephone as well as face-to-face appointments. All TEA participants received physiotherapy that included assessment and reassessment (100%) and most received physiotherapy that included education and advice (99%) and supervised exercises in clinic (94%) that included lower limb/knee exercises (99%) and prescribed general exercise (76%). By far, the most commonly recommended general physical activity was walking, or combinations of walking and cycling or walking and swimming. However, there was evidence from the CRFs for the TEA intervention of individualisation of the prescribed activities to participants' preferences including, bowling, going to the gym, running, golf, Zumba® (Zumba Fitness, LLC, Hallandale, FL, USA), dancing, pilates, football and tennis. Of the adherence-enhancing toolkit components, the most commonly used were the educational strategies: written educational material for patients (86%) and the individualised PhysioTools exercise sheets (85%). The behavioural strategies were used for, at most, two-thirds of patients only; the most commonly used were the physical activity diary (67%) followed by pedometers (for 53% of participants) and graded activity charts (15%). Heart rate monitoring during general exercise, visual feedback charts and reminder postcards were used with < 10 participants in total. Physiotherapists used the cognitive–behavioural strategies available to them within the toolkit as follows:

- questions to elicit health-related beliefs (36%)
- set-back plans with patients (36%)
- assessment of patient barriers and motivators to physical activity (35%)
- SMART goal-setting/contracting (30%)
- readiness/importance/confidence rulers (13%)
- decisional balance sheets (1%).

A clear signpost to local exercise opportunities/facilities in the geographical locality of the participant was recorded on the CRFs for 52% of those randomised to the TEA intervention arm. In total, 79 (48%) of the CRFs were judged to be per protocol for the TEA intervention.

Treatments other than exercise formed part of the package of care for 23%, 33% and 23% of participants in UC, ITE and TEA, respectively. The three most common additional treatments were advice about orthotics, ice therapy and manual therapy.

No SAEs and four adverse events were reported attributable to the interventions: one in UC (a sprained ankle), two in ITE (a sprained ankle and a twisted painful knee) and one in TEA (a fall during walking). Out of 514 participants, expected muscle soreness or transient increases in pain or aching occurred in 82 participants (19% in UC, 20% in ITE and 12% in TEA). Over the course of the 18 months of follow-up, there was one death in the BEEP trial (notified July 2012), which was deemed not attributable to the BEEP trial interventions.

Clinical outcomes (primary and secondary)

Primary outcome data were obtained from 87% of participants at 6 months (89% in UC, 86% in ITE and 85% in TEA), 77% at 9 months (79% in UC, 76% in ITE and 75% in TEA) and 79% at 18 months (81% in UC, 80% in ITE and 76% in TEA). Participants lost to follow-up at 6 months had slightly worse baseline knee pain stiffness and function scores (as measured by the WOMAC) and slightly higher levels of anxiety or depression at baseline than those who did not respond. They also were less likely to have used local facilities or opportunities that involved any form of physical activity in the last 7 days. (Data not shown but available from authors on request.)

The primary analyses of the trial showed that there were no statistically significant differences in the change in WOMAC pain or function at 6 months between UC and either of the other physiotherapy-led interventions (*Figure 9* and *Table 35*). Participants had WOMAC pain scores of 8.2 in the UC arm, 8.5 in the ITE arm and 8.5 in the TEA intervention arm at baseline and these reduced to 6.6, 6.8 and 6.8 at 3 months, respectively, and 6.4, 6.4 and 6.2 at 6 months, respectively, with little further change through to 18 months' follow-up (6.2, 6.1 and 6.1, respectively). In terms of WOMAC function, baseline scores were 27.5 in the UC arm, 27.8 in the ITE arm and 29.1 in the TEA arm, which reduced by approximately 4 points in all three groups by 3 months (23.2, 23.1 and 24.7, respectively) and between 5 and 8 WOMAC points by 6 months (21.4, 22.3 and 21.5, respectively). Longer-term outcomes at 9 and 18 months can be seen in *Table 35* but remained at similar levels to those seen at 6 months. A longitudinal analysis of the mean outcome trajectory for the primary outcomes also did not show any statistically significant differences in the mean trajectory by treatment arm (data not shown). In terms of effect size estimates for each intervention arm (comparing follow-up scores with baseline scores), all three interventions led to improvements of moderate effect size at both 6 months (WOMAC pain Cohen's effect size estimates of UC 0.56, ITE 0.57, TEA 0.66; and WOMAC function UC 0.51, ITE 0.43, TEA 0.63) and 18 months (WOMAC pain UC 0.63, ITE 0.65, TEA 0.69; and WOMAC function UC 0.50, ITE 0.46, TEA 0.51). However, these should be interpreted cautiously given the lack of a no treatment control group in the comparison.

The sensitivity analyses did not change the interpretation of the trial results. The per protocol analysis (including only those participants who were judged to have received the BEEP trial interventions in line with the treatment protocols) also demonstrated no statistically significant differences in the change in WOMAC pain or function scores, although *Table 35* shows a trend for lower pain and function scores in those in the TEA intervention group who received the intervention per protocol. Results for the primary outcome at the primary end point also did not differ depending on whether the participant showed high or low levels of self-reported exercise adherence (*Table 36*) or when complete case data were analysed, when baseline covariates were excluded from the analysis and when any group effect from therapists working in the same clinic ('clustering') has been accounted for in the treatment models (data not shown).

In terms of secondary outcomes (*Table 37*), the analyses showed consistent results overall, of no statistically significant differences in the change in outcomes between UC and either the ITE intervention group or the TEA intervention group. In total, at 6 months, 50% of those in the UC group could be classified as treatment responders using the OARSI responder criteria, compared with 51% in the ITE group and 55% in the TEA intervention group (differences that were not statistically significant). Interestingly, these proportions remained relatively stable over longer-term follow-ups (48%, 50% and 51%, respectively, at 18-month follow-up).

Self-reported physical activity, measured using the PASE, was not statistically significantly different between the intervention groups and the mean trajectory for this outcome over time showed no significant differences between treatment arms. Participants had scores of 176 in the UC group, 175 in the ITE group and 180 in the TEA group at baseline and these increased by similar amounts in all three groups by 6 months (188, 189 and 196, respectively), indicating that all three groups reported small increases in their general physical activity levels. However, by the 18-month follow-up, these self-reported physical activity levels had returned to baseline (or below baseline) levels (177, 173 and 169, respectively). Participants' self-reported use of physical activity facilities in the last 7 days appeared to increase in all three treatment

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

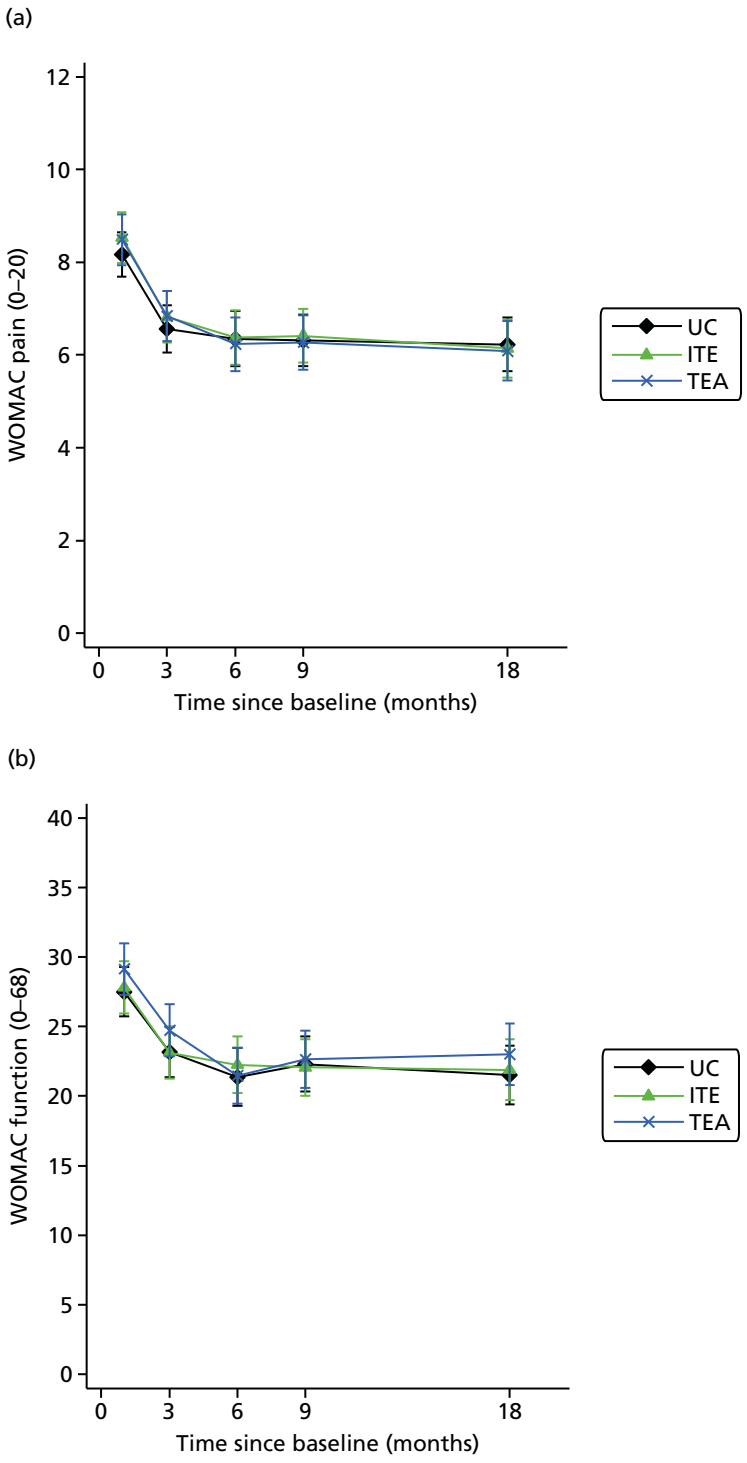

FIGURE 9 Means and 95% CIs for the primary outcome at each time point. (a) WOMAC pain; and (b) WOMAC function.

groups (increasing from 36% in UC, 32% in ITE and 29% in TEA at baseline to 51%, 43% and 49%, respectively, at 6 months) and these proportions appeared to stay high even at longer-term follow-up, but there was no further increase. The OR at 9 months in the TEA group compared with UC, for a report of using local physical activity facilities in the last 7 days, was 1.7 (95% CI 1.0 to 3.0) but there were no differences between the groups at 18 months.

TABLE 35 Summary of results for primary outcomes at each time point

Outcome measure	3 months	6 months	9 months	18 months
***Imputed data:* N**	***514***	***514***	***514***	***514***
WOMAC pain (0–20)				
UC: mean (SD)	6.6 (3.4)	6.4 (4.0)	6.3 (3.7)	6.2 (3.9)
ITE: mean (SD)	6.8 (3.7)	6.4 (4.0)	6.4 (3.9)	6.1 (4.2)
TEA: mean (SD)	6.8 (3.5)	6.2 (3.8)	6.3 (3.8)	6.1 (4.2)
UC vs. ITE: adjusted mean difference (95% CI)	0.0 (–0.6 to 0.6)	–0.3 (–1.0 to 0.5)	–0.2 (–0.9 to 0.6)	–0.3 (–1.2 to 0.5)
UC vs. TEA: adjusted mean difference (95% CI)	0.1 (–0.6 to 0.7)	–0.3 (–1.0 to 0.4)	–0.3 (–1.0 to 0.5)	–0.3 (–1.2 to 0.5)
WOMAC function (0–68)				
UC: mean (SD)	23.2 (12.3)	21.4 (14.1)	22.3 (13.3)	21.5 (14.4)
ITE: mean (SD)	23.1 (12.7)	22.3 (13.7)	22.1 (13.8)	21.9 (14.9)
TEA: mean (SD)	24.7 (12.3)	21.5 (13.2)	22.7 (13.3)	23.0 (14.4)
UC vs. ITE: adjusted mean difference (95% CI)	–0.4 (–2.5 to 1.8)	0.4 (–2.0 to 2.8)	–0.6 (–3.2 to 1.9)	–0.2 (–3.0 to 2.6)
UC vs. TEA: adjusted mean difference (95% CI)	0.5 (–1.7 to 2.6)	–1.0 (–3.4 to 1.4)	–0.7 (–3.3 to 1.9)	0.4 (–2.6 to 3.3)
Per-protocol analysis (imputed data)				
WOMAC pain (0–20)				
UC: *n*, mean (SD)	156, 6.4 (3.5)	156, 6.2 (4.1)	156, 6.2 (3.8)	156, 6.1 (3.9)
ITE: *n*, mean (SD)	109, 6.8 (3.6)	109, 6.4 (4.0)	109, 6.5 (3.9)	109, 6.2 (4.1)
TEA: *n*, mean (SD)	78, 6.3 (2.8)	78, 5.8 (3.4)	78, 5.9 (3.7)	78, 5.3 (3.7)
UC vs. ITE: adjusted mean difference (95% CI)	0.1 (–0.6 to 0.8)	–0.2 (–1.0 to 0.6)	–0.1 (–0.9 to 0.8)	–0.2 (–1.1 to 0.7)
UC vs. TEA: adjusted mean difference (95% CI)	–0.3 (–1.0 to 0.5)	–0.6 (–1.5 to 0.3)	–0.4 (–1.4 to 0.5)	–0.8 (–1.8 to 0.2)
WOMAC function (0–68)				
UC: *n*, mean (SD)	156, 22.5 (12.4)	156, 20.5 (14.1)	156, 22.0 (13.4)	156, 20.9 (14.3)
ITE: *n*, mean (SD)	109, 22.6 (12.6)	109, 21.7 (13.8)	109, 22.1 (13.7)	109, 21.6 (14.5)
TEA: *n*, mean (SD)	78, 22.6 (10.0)	78, 20.5 (11.8)	78, 20.8 (12.6)	78, 20.4 (12.7)
UC vs. ITE: adjusted mean difference (95% CI)	–0.5 (–2.8 to 1.8)	0.1 (–2.6 to 2.9)	–0.6 (–3.6 to 2.3)	–0.2 (–3.3 to 2.9)
UC vs. TEA: adjusted mean difference (95% CI)	–1.2 (–3.8 to 1.5)	–1.4 (–4.5 to 1.6)	–2.2 (–5.5 to 1.0)	–1.3 (–4.7 to 2.0)

Figures are presented after imputation of missing data and after adjustment for baseline WOMAC scores, age, sex, onset of knee problem, and treatment centre unless otherwise stated. Higher WOMAC scores = worse outcome.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 36 Analysis of the primary outcome at the primary end point (i.e. the 6-month follow-up) stratified by level of self-reported exercise adherence at 6-month follow-up

	I have been doing my exercises as often as I was advised to . . .	
	Strongly agree/agree (*N* = 268)	Strongly disagree/disagree (*N* = 84)
WOMAC pain (0–20)		
UC: mean (SD)	6.4 (4.0)	6.1 (3.8)
ITE: mean (SD)	6.1 (4.2)	6.2 (3.5)
TEA: mean (SD)	5.9 (3.8)	6.2 (3.7)
Interaction regression coefficient[a] (95% CI)		
UC	0	0
ITE	0	1.0 (–0.9, 2.9)
TEA	0	0.6 (–1.4, 2.6)
WOMAC function (0–68)		
UC: mean (SD)	22.1 (14.0)	20.0 (13.9)
ITE: mean (SD)	20.9 (14.7)	21.0 (12.4)
TEA: mean (SD)	20.8 (13.2)	19.6 (14.1)
Interaction regression coefficient[a] (95% CI)		
UC	0	0
ITE	0	4.8 (–1.7, 11.2)
TEA	0	1.5 (–5.3, 8.3)

a The interaction regression coefficients express the interaction between treatment arm and level of exercise adherence after the outcome of interest has been adjusted in a regression model for baseline in the outcome of interest, age, sex, duration of the knee problem and physiotherapy treatment centre (as was used in the randomisation algorithm).
Fifty-three participants reported they were 'not sure' if they had been doing their exercises as often as advised so are excluded from the analysis.

TABLE 37 Summary of results for all secondary outcome measures

Outcome measure	3 months (*N* = 514)	6 months (*N* = 514)	9 months (*N* = 514)	18 months (*N* = 514)
OARSI responder criteria[a]				
UC: *n* (%)	77 (44)	88 (50)	79 (45)	84 (48)
ITE: *n* (%)	81 (46)	90 (51)	86 (49)	88 (50)
TEA: *n* (%)	73 (45)	90 (55)	82 (50)	83 (51)
UC vs. ITE: adjusted OR (95% CI)	1.1 (0.7 to 1.8)	1.1 (0.7 to 1.8)	1.3 (0.8 to 2.0)	1.1 (0.7 to 1.9)
UC vs. TEA: adjusted OR (95% CI)	1.0 (0.6 to 1.7)	1.2 (0.8 to 2.0)	1.3 (0.8 to 2.1)	1.1 (0.7 to 1.9)
WOMAC stiffness (0–8)				
UC: mean (SD)	3.1 (1.8)	2.9 (1.8)	3.0 (1.8)	2.8 (1.8)
ITE: mean (SD)	3.1 (1.7)	3.1 (1.9)	3.0 (1.8)	2.8 (1.8)
TEA: mean (SD)	3.1 (1.7)	2.9 (1.7)	3.1 (1.7)	3.1 (1.9)
UC vs. ITE: adjusted mean difference (95% CI)	0.0 (–0.3 to 0.3)	0.2 (–0.2 to 0.5)	0.0 (–0.4 to 0.3)	0.0 (–0.3 to 0.4)
UC vs. TEA: adjusted mean difference (95% CI)	0.0 (–0.4 to 0.3)	–0.1 (–0.4 to 0.3)	0.1 (–0.3 to 0.4)	0.3 (–0.1 to 0.6)

TABLE 37 Summary of results for all secondary outcome measures (*continued*)

Outcome measure	3 months (*N* = 514)	6 months (*N* = 514)	9 months (*N* = 514)	18 months (*N* = 514)
PASE (0–590)				
UC: mean (SD)	196 (87)	188 (84)	170 (77)	177 (89)
ITE: mean (SD)	188 (86)	189 (89)	162 (85)	173 (80)
TEA: mean (SD)	193 (92)	196 (95)	187 (88)	169 (78)
UC vs. ITE: adjusted mean difference (95% CI)	–6.4 (–26.1 to 13.3)	3.8 (–14.8 to 22.4)	–4.5 (–23.6 to 14.6)	–2.3 (–20.8 to 16.3)
UC vs. TEA: adjusted mean difference (95% CI)	–4.1 (–23.4 to 15.2)	8.3 (–10.7 to 27.3)	18.7 (–1.3 to 38.6)	–7.3 (–25.8 to 11.3)
BMI (kg/m^2)				
UC: mean (SD)	29.2 (5.7)	29.2 (5.6)	29.2 (5.7)	28.9 (5.4)
ITE: mean (SD)	29.3 (5.6)	29.1 (5.8)	29.3 (5.8)	28.8 (5.6)
TEA: mean (SD)	30.0 (5.5)	29.7 (5.5)	29.7 (5.5)	29.7 (5.4)
UC vs. ITE: adjusted mean difference (95% CI)	0.1 (–0.2 to 0.4)	0.0 (–0.4 to 0.3)	0.2 (–0.3 to 0.6)	–0.1 (–0.6 to 0.4)
UC vs. TEA: adjusted mean difference (95% CI)	0.3 (0.0 to 0.6)	0.0 (–0.3 to 0.4)	0.0 (–0.5 to 0.5)	0.3 (–0.1 to 0.8)
Used local facilities or opportunities that involved any form of physical activity in the last 7 days				
UC: *n* (%)	100 (57)	89 (51)	75 (43)	82 (47)
ITE: *n* (%)	95 (54)	76 (43)	65 (37)	90 (51)
TEA: *n* (%)	88 (54)	80 (49)	82 (50)	70 (43)
UC vs. ITE: adjusted OR (95% CI)	0.9 (0.6 to 1.5)	0.8 (0.5 to 1.3)	0.9 (0.5 to 1.5)	1.3 (0.8 to 2.3)
UC vs. TEA: adjusted OR (95% CI)	1.1 (0.6 to 1.8)	1.1 (0.6 to 1.8)	1.7 (1.0 to 3.0)	1.0 (0.6 to 1.7)
GAD-7 (0–21)				
UC: median (IQR)	1.0 (0.0–3.4)	1.0 (0.0–4.0)	[b]	1.0 (0.0–4.2)
ITE: median (IQR)	1.0 (0.0–4.0)	1.0 (0.0–3.8)	[b]	1.0 (0.0–4.6)
TEA: median (IQR)	1.0 (0.0–4.6)	1.0 (0.0–3.9)	[b]	0.5 (0.0–3.8)
UC vs. ITE: adjusted mean[c] difference (95% CI)	0.2 (–0.5 to 0.9)	–0.1 (–0.9 to 0.6)	[b]	0.4 (–0.4 to 1.2)
UC vs. TEA: adjusted mean[c] difference (95% CI)	0.5 (–0.2 to 1.2)	0.2 (–0.6 to 1.0)	[b]	0.0 (–0.9 to 0.8)
PHQ-8 depressive scale (0–24)				
UC: median (IQR)	2.0 (0.2–5.0)	2.0 (0.0–4.0)	[b]	2.0 (0.0–6.1)
ITE: median (IQR)	2.0 (0.0–5.0)	2.0 (0.0–4.4)	[b]	2.3 (0.0–6.0)
TEA: median (IQR)	2.0 (0.0–5.1)	1.6 (0.0,4.2)	[b]	2.1 (0.0–5.2)
UC vs. ITE: adjusted mean[c] difference (95% CI)	0.3 (–0.4 to 1.1)	0.4 (–0.4 to 1.3)	[b]	0.5 (–0.4 to 1.4)
UC vs. TEA: adjusted mean[c] difference (95% CI)	0.5 (–0.3 to 1.3)	0.6 (–0.3 to 1.4)	[b]	0.3 (–0.6 to 1.3)
SEE scale (0–10)				
UC: mean (SD)	5.6 (2.1)	5.4 (2.3)	[b]	[b]
ITE: mean (SD)	5.9 (2.3)	5.7 (2.1)	[b]	[b]
TEA: mean (SD)	5.5 (2.4)	5.7 (2.2)	[b]	[b]

continued

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 37 Summary of results for all secondary outcome measures (*continued*)

Outcome measure	3 months (*N* = 514)	6 months (*N* = 514)	9 months (*N* = 514)	18 months (*N* = 514)
UC vs. ITE: adjusted mean difference (95% CI)	0.3 (–0.2 to 0.8)	0.3 (–0.1 to 0.8)	[b]	[b]
UC vs. TEA: adjusted mean difference (95% CI)	–0.1 (–0.5 to 0.4)	0.4 (–0.2 to 0.8)	[b]	[b]
OEE scale – positive subscale (1–5)				
UC: mean (SD)	4.0 (0.6)	4.0 (0.6)	[b]	[b]
ITE: mean (SD)	4.0 (0.6)	4.0 (0.6)	[b]	[b]
TEA: mean (SD)	3.9 (0.5)	3.9 (0.6)	[b]	[b]
UC vs. ITE: adjusted mean difference (95% CI)	0.1 (–0.1 to 0.2)	0.0 (–0.1 to 0.1)	[b]	[b]
UC vs. TEA: adjusted mean difference (95% CI)	–0.1 (–0.2 to 0.1)	0.0 (–0.2 to 0.1)	[b]	[b]
OEE scale – negative subscale (1–5)[d]				
UC: mean (SD)	3.8 (0.7)	3.8 (0.8)	[b]	[b]
ITE: mean (SD)	3.9 (0.8)	3.9 (0.9)	[b]	[b]
TEA: mean (SD)	3.7 (0.8)	3.8 (0.8)	[b]	[b]
UC vs. ITE: adjusted mean difference (95% CI)	0.1 (–0.1 to 0.2)	0.1 (–0.1 to 0.3)	[b]	[b]
UC vs. TEA: adjusted mean difference (95% CI)	0.0 (–0.2 to 0.1)	0.2 (0.0 to 0.3)	[b]	[b]
IPQR – How much does your knee pain affect your life? (0–10)				
UC: mean (SD)	4.5 (2.4)	4.3 (2.5)	[b]	[b]
ITE: mean (SD)	4.4 (2.3)	4.2 (2.5)	[b]	[b]
TEA: mean (SD)	4.6 (2.4)	4.3 (2.5)	[b]	[b]
UC vs. ITE: adjusted mean difference (95% CI)	–0.1 (–0.6 to 0.3)	–0.1 (–0.6 to 0.4)	[b]	[b]
UC vs. TEA: adjusted mean difference (95% CI)	0.1 (–0.3 to 0.6)	0.0 (–0.5 to 0.5)	[b]	[b]
IPQR – How long do you think your knee pain will continue? (0–10)				
UC: median (IQR)	8.0 (5.1–10.0)	8.5 (5.5–10.0)	[b]	[b]
ITE: median (IQR)	8.0 (5.0–10.0)	8.0 (5.1–10.0)	[b]	[b]
TEA: median (IQR)	7.8 (5.0–10.0)	8.0 (5.0–10.0)	[b]	[b]
UC vs. ITE: adjusted mean[c] difference (95% CI)	–0.1 (–0.7 to 0.5)	–0.1 (–0.7 to 0.6)	[b]	[b]
UC vs. TEA: adjusted mean[c] difference (95% CI)	–0.2 (–0.8 to 0.4)	–0.2 (–0.9 to 0.4)	[b]	[b]
IPQR – How much control do you feel you have over your knee pain? (0–10)				
UC: mean (SD)	5.0 (2.7)	4.9 (2.5)	[b]	[b]
ITE: mean (SD)	5.2 (2.6)	5.4 (2.6)	[b]	[b]
TEA: mean (SD)	4.6 (2.6)	5.3 (2.6)	[b]	[b]
UC vs. ITE: adjusted mean difference (95% CI)	0.3 (–0.4 to 0.9)	0.5 (–0.1 to 1.1)	[b]	[b]
UC vs. TEA: adjusted mean difference (95% CI)	–0.3 (–0.9 to 0.3)	0.4 (–0.2 to 1.1)	[b]	[b]

TABLE 37 Summary of results for all secondary outcome measures (*continued*)

Outcome measure	3 months (*N* = 514)	6 months (*N* = 514)	9 months (*N* = 514)	18 months (*N* = 514)
IPQR – How much do you think treatment can help your knee pain? (0–10)				
UC: median (IQR)	6.1 (5.0–8.0)	6.0 (4.2–8.0)	[b]	[b]
ITE: median (IQR)	7.0 (5.0–8.0)	7.0 (4.2–8.0)	[b]	[b]
TEA: median (IQR)	7.0 (5.0–8.0)	7.0 (5.0–9.0)	[b]	[b]
UC vs. ITE: adjusted mean[c] difference (95% CI)	0.2 (–0.3 to 0.8)	0.3 (–0.3 to 0.9)	[b]	[b]
UC vs. TEA: adjusted mean[c] difference (95% CI)	0.1 (–0.4 to 0.7)	0.7 (0.1 to 1.4)	[b]	[b]
IPQR – How much do you experience symptoms from your knee pain? (0–10)				
UC: mean (SD)	5.3 (2.3)	5.0 (2.3)	[b]	[b]
ITE: mean (SD)	5.2 (2.3)	4.9 (2.5)	[b]	[b]
TEA: mean (SD)	5.4 (2.3)	5.0 (2.3)	[b]	[b]
UC vs. ITE: adjusted mean difference (95% CI)	–0.1 (–0.6 to 0.4)	–0.1 (–0.6 to 0.4)	[b]	[b]
UC vs. TEA: adjusted mean difference (95% CI)	0.2 (–0.3 to 0.6)	0.0 (–0.5 to 0.5)	[b]	[b]
IPQR – How concerned are you about your knee pain? (0–10)				
UC: median (IQR)	6.6 (5.0–8.0)	5.3 (3.0–7.9)	[b]	[b]
ITE: median (IQR)	6.1 (3.9–8.4)	5.5 (3.0–8.0)	[b]	[b]
TEA: median (IQR)	6.4 (5.0–8.1)	5.1 (3.0–8.0)	[b]	[b]
UC vs. ITE: adjusted mean[c] difference (95% CI)	0.0 (–0.6 to 0.6)	0.2 (–0.4 to 0.8)	[b]	[b]
UC vs. TEA: adjusted mean[c] difference (95% CI)	0.1 (–0.4 to 0.7)	0.1 (–0.6 to 0.7)	[b]	[b]
IPQR – How well do you feel you understand your knee pain? (0–10)				
UC: mean (SD)	6.9 (2.5)	7.1 (2.4)	[b]	[b]
ITE: mean (SD)	7.2 (2.5)	7.6 (2.5)	[b]	[b]
TEA: mean (SD)	6.8 (2.6)	7.5 (2.5)	[b]	[b]
UC vs. ITE: adjusted mean difference (95% CI)	0.2 (–0.3 to 0.8)	0.4 (–0.1 to 1.0)	[b]	[b]
UC vs. TEA: adjusted mean difference (95% CI)	0.0 (–0.5 to 0.6)	0.5 (0.0 to 1.1)	[b]	[b]
IPQR – How much does your knee pain affect you emotionally? (0–10)				
UC: mean (SD)	4.1 (2.8)	3.4 (2.9)	[b]	[b]
ITE: mean (SD)	3.8 (2.9)	3.6 (2.9)	[b]	[b]
TEA: mean (SD)	4.3 (2.8)	3.6 (2.9)	[b]	[b]
UC vs. ITE: adjusted mean difference (95% CI)	–0.3 (–0.9 to 0.2)	0.1 (–0.5 to 0.7)	[b]	[b]
UC vs. TEA: adjusted mean difference (95% CI)	0.1 (–0.5 to 0.6)	0.2 (–0.5 to 0.8)	[b]	[b]

continued

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 37 Summary of results for all secondary outcome measures (*continued*)

Outcome measure	3 months (*N* = 514)	6 months (*N* = 514)	9 months (*N* = 514)	18 months (*N* = 514)
Accelerometer data[e] (subsample of trial participants)	**N = *50***	**N = *47***	**N = *51***	**N = *48***
Average number of counts per minute (measured by accelerometers)				
UC: *n*: median (IQR)	15–346 (124–460)	14–338 (138–424)	15–306 (176–414)	14–222 (117–440)
ITE: *n*: median (IQR)	19–257 (170–278)	17–209 (159–290)	20–201 (157–316)	18–238 (174–348)
TEA: *n*: median (IQR)	16–239 (193–324)	16–264 (205–278)	16–223 (166–250)	16–205 (161–283)
UC vs. ITE: adjusted[f] mean[c] difference (95% CI)	–28 (–84 to 28)	–66 (–147 to 15)	–47 (–101 to 7)	–17 (–86 to 52)
UC vs. TEA: adjusted[f] mean[c] difference (95% CI)	–51 (–108 to 7)	–53 (–138 to 31)	–61 (–117 to -4)	–56 (–130 to 19)
Proportion meeting physical activity guidelines (measured by accelerometers[g])				
UC: *n* (%)	5 (33)	4 (29)	5 (33)	3 (21)
ITE: *n* (%)	0 (0)	0 (0)	0 (0)	3 (17)
TEA: *n* (%)	3 (19)	1 (6)	0 (0)	1 (6)
UC vs. ITE: adjusted[f] OR (95% CI)	[h]	[h]	[h]	[h]
UC vs. TEA: adjusted[f] OR (95% CI)	[h]	[h]	[h]	[h]

IPQR, Brief Illness Perceptions Questionnaire; PHQ-8, Patient Health Questionnaire 8.
Figures are presented after imputation of missing data (with the exception of data from the accelerometers) and after adjustment for the baseline score on the outcome of interest (with the exception of the OARSI responder criteria), age, sex, onset of knee problem, and treatment centre, unless otherwise stated.
WOMAC stiffness higher score = more severe stiffness; PASE higher score = more active; SEE higher score = more confident that exercise can be done; OEE positive and negative subscales higher score = higher expectations that exercise will be beneficial; GAD-7 higher score = more anxious; PHQ-8 higher score = more depressed; IPQR – affects life, higher score = more affected; IPQR – duration, higher score = lasts a longer time; IPQR – personal control, higher score = more control; IPQR – treatment control, higher score = higher belief treatment can control; IPQR – symptom experience, higher score = more symptoms that are more severe; IPQR – concern, higher score = more concerned; IPQR – understanding, higher score = more understanding; IPQR – emotion, higher score = more emotionally affected.

a Participants met the OARSI responder criteria if (1) relative change in WOMAC pain or function was ≥ 50% and absolute change was ≥ 20 or (2) at least two of the following applied: relative change in pain ≥ 20% and absolute change ≥ 10, relative change in function ≥ 20% and absolute change ≥ 10 or participants reported they were better, much better, or completely recovered on the global assessment of change question. Absolute change (baseline – follow-up score) and relative change (absolute change/baseline score) were calculated after WOMAC measures were scaled from 1 to 101 to avoid dividing by 0 when calculating relative change (Pham *et al.*[201]).
b Data not collected at this time point.
c Mean differences are presented, despite a skewed distribution for the outcome at the absolute time point, as, when adjusted for the baseline value of interest, model residuals followed a normal distribution.
d Reverse scored (i.e. a higher score on the negative subscale indicates higher expectations of the benefit of exercise).
e Participants are only included if they have worn the monitor for at least 5 days for ≥ 10 hours. Valid time is calculated assuming that any consecutive runs of zero count lasting for ≥ 60 minutes are counted as non-wear.
f Model adjusted for baseline only (adjusting for all a priori model covariates gave unstable model results owing to the small sample size used for the analysis).
g Defined as participants completing 150 minutes each week of moderate intensity physical activity (accumulated in bouts of ≥ 10 minutes) or 75 minutes of vigorous intensity activity spread across the week (adapted from Dh *Start Active, Stay Active: A Report on Physical Activity for Health from the Four Home Countries' Chief Medical Officers*[242]). Bouts calculated using a drop-time of 2 minutes (ActiGraph Support[250]). Missing days of data are imputed using the average of the average count for days for which data are present.
h Not calculated owing to small *n*.

The objective data on actual physical activity levels, measured in the subsample of participants allocated to wear accelerometers at each follow-up time point, showed somewhat different results. Although, again, there were no statistically significant differences in the changes in activity counts between UC and the other two intervention groups, the trends were consistently in the direction of higher physical activity levels in the UC group than in either the ITE or TEA intervention groups. In both the UC and ITE groups, objective physical activity counts increased during the period of physiotherapy treatment (from baseline to 3 months) and then reduced to below baseline levels by 18 months. In the TEA group, for which increasing general physical activity was a key focus of the intervention, objective physical activity counts increased during the period of physiotherapy treatment (from baseline to 6 months) but again fell to below baseline levels by the 18-month follow-up. The proportions of participants who could be classified as meeting recommended physical activity levels, based on their accelerometer data, were very small but seemed to be highest in the UC group (see *Table 37*). In addition, when the data were categorised into pre-defined levels of activity, it was found that the majority of participants' time was spent in sedentary activity, irrespective of treatment arm or time point considered (data not shown but available from authors on request).

In terms of participants' illness perceptions (measured at baseline and 3 and 6 months only), there were very few differences between the intervention groups. One exception was participants' views on how much treatment could help their knee pain, which appeared to be statistically significantly greater at 6 months in the TEA group than in the UC group (see *Table 37*). In general, participants in all three groups appeared to benefit from the BEEP trial interventions over time, with all groups reporting that they felt they had greater control over their knee problem and were less concerned about their knee problem than at baseline.

Intervention credibility and exercise adherence outcomes

Table 38 summarises the results for treatment credibility and exercise adherence at each follow-up time point. Overall, treatment credibility was high in all groups and remained so even at the 18-month follow-up. The majority of participants who received UC, ITE and TEA felt that their treatment was quite or very logical, and were quite or very confident that (1) the treatment they received could help their knee problem, (2) they could recommend the treatment they received to a friend and (3) the treatment would be successful in helping other types of problems.

At 3 months' follow-up, self-reported exercise adherence was high in all groups, with ≥ 75% agreeing or strongly agreeing that they had completed their exercises as often as they had been advised. In the UC group this reduced to 55% at 6 months and then fell to < 50% at longer-term follow-ups. In the ITE group, adherence levels also steadily reduced over time, although to a lesser extent than in the UC group, with adherence dropping only < 50% by 18 months' follow-up. Adherence to exercise in participants randomised to the TEA intervention remained consistently high until the 9-month follow-up (6 months: 77%; 9 months: 74%). At 18 months, although adherence was higher in the TEA group than in the other two intervention groups, the proportion of participants agreeing or strongly agreeing that they had been doing their exercises as often as advised had reduced to 51%.

At 3 months, more than half of those in the ITE and TEA groups reported exercising every day in comparison with the UC group (in which more than half of participants reported exercising five times per week). At 6 months, more than half of those in the TEA group reported exercising six times per week in comparison with four times per week in the ITE and UC intervention arms. Similarly, at 9 and 18 months, those in the TEA group reported exercising more frequently than those in the other two groups (as judged by the median response category reported). There were no differences in the exercise durations between the intervention groups.

Health economic outcomes

Parts of this section have been reproduced from Kigozi *et al.*[265] © The Author(s) 2018. Published by Oxford University Press on behalf of the British Society for Rheumatology. This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 38 Treatment credibility and exercise adherence at each follow-up time point

Question	3 months			6 months			9 months			18 months		
	UC (*n* = 143)	ITE (*n* = 146)	TEA (*n* = 136)	UC (*n* = 142)	ITE (*n* = 140)	TEA (*n* = 131)	UC (*n* = 139)	ITE (*n* = 135)	TEA (*n* = 129)	UC (*n* = 141)	ITE (*n* = 134)	TEA (*n* = 120)
Confident that treatment received can help knee problem												
Very confident	30 (21)	49 (34)	45 (33)	26 (19)	43 (31)	51 (40)	22 (16)	31 (23)	39 (31)	15 (11)	30 (23)	31 (27)
Quite confident	70 (50)	66 (46)	63 (47)	63 (45)	62 (45)	45 (35)	59 (44)	66 (49)	53 (42)	59 (44)	59 (44)	49 (42)
Neither	13 (9)	9 (6)	11 (8)	15 (11)	8 (6)	13 (10)	23 (17)	13 (10)	9 (7)	28 (21)	15 (11)	13 (11)
Not very confident	22 (16)	16 (11)	14 (10)	26 (19)	19 (14)	10 (8)	24 (18)	17 (13)	18 (14)	24 (18)	19 (14)	15 (13)
Not at all confident	5 (4)	5 (3)	2 (1)	9 (6)	6 (4)	10 (8)	6 (4)	7 (5)	7 (6)	9 (7)	10 (8)	9 (8)
Confident in recommending this treatment to a friend												
Very confident	33 (24)	57 (39)	50 (37)	32 (23)	59 (43)	50 (39)	29 (21)	48 (36)	46 (37)	26 (19)	39 (29)	41 (35)
Quite confident	70 (50)	59 (41)	60 (45)	63 (45)	47 (34)	51 (40)	58 (43)	57 (43)	50 (40)	61 (45)	54 (41)	45 (38)
Neither	16 (11)	13 (9)	13 (10)	22 (16)	21 (15)	9 (7)	31 (23)	14 (10)	10 (8)	26 (19)	13 (10)	15 (13)
Not very confident	19 (14)	10 (7)	8 (6)	17 (12)	7 (5)	12 (9)	10 (7)	10 (7)	16 (13)	18 (13)	15 (11)	10 (9)
Not at all confident	2 (1)	6 (4)	3 (2)	6 (4)	4 (3)	7 (5)	7 (5)	5 (4)	4 (3)	4 (3)	12 (9)	6 (5)
Treatment makes sense to you												
Very logical	46 (33)	73 (50)	57 (42)	38 (27)	70 (51)	61 (47)	39 (29)	51 (38)	48 (38)	36 (27)	45 (34)	40 (34)
Quite logical	76 (54)	61 (42)	68 (50)	78 (56)	50 (36)	54 (42)	70 (52)	65 (49)	54 (43)	64 (47)	65 (49)	51 (44)
No opinion	12 (9)	3 (2)	6 (4)	17 (12)	9 (7)	6 (5)	21 (16)	8 (6)	11 (9)	26 (19)	14 (11)	14 (12)
Not very logical	5 (4)	6 (4)	3 (2)	7 (5)	7 (5)	5 (4)	5 (4)	7 (5)	11 (9)	8 (6)	2 (2)	3 (3)
Not at all logical	1 (1)	2 (1)	1 (1)	0 (0)	2 (1)	3 (2)	0 (0)	3 (2)	2 (2)	1 (1)	7 (5)	9 (8)

Question	3 months			6 months			9 months			18 months		
	UC (*n* = 143)	ITE (*n* = 146)	TEA (*n* = 136)	UC (*n* = 142)	ITE (*n* = 140)	TEA (*n* = 131)	UC (*n* = 139)	ITE (*n* = 135)	TEA (*n* = 129)	UC (*n* = 141)	ITE (*n* = 134)	TEA (*n* = 120)
Treatment would be successful in helping other types of problems												
Very successful	23 (16)	22 (15)	30 (22)	18 (13)	27 (20)	39 (30)	24 (18)	23 (17)	30 (24)	21 (16)	26 (20)	23 (20)
Quite successful	55 (39)	73 (50)	62 (46)	65 (46)	60 (43)	43 (33)	51 (38)	71 (53)	43 (34)	52 (39)	56 (42)	50 (43)
No opinion	54 (39)	46 (32)	41 (30)	54 (39)	46 (33)	40 (31)	53 (39)	34 (25)	46 (36)	52 (39)	43 (33)	35 (30)
Not very successful	7 (5)	2 (1)	1 (1)	3 (2)	4 (3)	6 (5)	6 (4)	3 (2)	7 (6)	9 (7)	5 (4)	4 (3)
Not at all successful	1 (1)	2 (1)	1 (1)	0 (0)	1 (1)	1 (1)	1 (1)	3 (2)	1 (1)	1 (1)	2 (2)	5 (4)
Been doing exercises as often as advised												
Strongly agree	23 (16)	39 (27)	39 (29)	12 (9)	27 (20)	36 (28)	15 (11)	23 (17)	22 (17)	12 (9)	19 (14)	18 (15)
Agree	82 (59)	78 (54)	71 (52)	64 (46)	66 (48)	63 (49)	48 (36)	52 (39)	72 (57)	50 (37)	42 (32)	42 (36)
Not sure	13 (9)	12 (8)	12 (9)	21 (15)	21 (15)	11 (9)	23 (17)	12 (9)	14 (11)	17 (13)	15 (11)	22 (19)
Disagree	20 (14)	13 (9)	12 (9)	34 (24)	20 (15)	17 (13)	36 (27)	39 (29)	17 (13)	42 (31)	44 (33)	25 (21)
Strongly disagree	2 (1)	3 (2)	2 (1)	9 (6)	3 (2)	1 (1)	11 (8)	7 (5)	2 (2)	13 (10)	12 (9)	10 (9)
Exercise frequency in the last month												
Never	6 (4)	4 (3)	2 (2)	18 (13)	11 (8)	3 (2)	21 (16)	20 (15)	16 (13)	36 (26)	34 (26)	30 (26)
Once a week	5 (4)	4 (3)	1 (1)	11 (8)	3 (2)	3 (2)	17 (13)	20 (15)	5 (4)	21 (15)	21 (16)	9 (8)
Twice a week	11 (8)	3 (2)	2 (2)	23 (17)	18 (13)	10 (8)	24 (18)	18 (14)	7 (6)	23 (17)	23 (18)	13 (11)
Three times a week	11 (8)	17 (12)	11 (8)	12 (9)	23 (17)	19 (15)	13 (10)	18 (14)	23 (18)	17 (13)	15 (11)	12 (10)
Four times a week	17 (12)	16 (11)	7 (5)	14 (10)	17 (12)	14 (11)	13 (10)	11 (8)	13 (10)	5 (4)	3 (2)	8 (7)
Five times a week	21 (15)	8 (6)	11 (8)	19 (14)	10 (7)	10 (8)	11 (8)	5 (4)	11 (9)	8 (6)	11 (8)	13 (11)
Six times a week	3 (2)	10 (7)	10 (8)	10 (7)	10 (7)	6 (5)	3 (2)	3 (2)	7 (6)	1 (1)	1 (1)	3 (3)
Once every day	47 (34)	54 (38)	43 (33)	22 (16)	32 (23)	30 (24)	26 (20)	34 (26)	30 (24)	20 (15)	19 (15)	21 (18)
Twice every day	19 (14)	27 (19)	44 (34)	8 (6)	13 (9)	31 (25)	4 (3)	2 (2)	15 (12)	5 (4)	4 (3)	6 (5)

continued

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 38 Treatment credibility and exercise adherence at each follow-up time point (*continued*)

Question	3 months			6 months			9 months			18 months		
	UC (*n* = 143)	ITE (*n* = 146)	TEA (*n* = 136)	UC (*n* = 142)	ITE (*n* = 140)	TEA (*n* = 131)	UC (*n* = 139)	ITE (*n* = 135)	TEA (*n* = 129)	UC (*n* = 141)	ITE (*n* = 134)	TEA (*n* = 120)
Exercise duration												
< 5 minutes	9 (6)	7 (5)	2 (2)	9 (7)	7 (5)	7 (5)	17 (13)	5 (4)	12 (10)	8 (6)	15 (12)	9 (8)
5 minutes to < 10 minutes	23 (17)	27 (19)	36 (27)	28 (21)	28 (20)	34 (27)	25 (20)	41 (32)	30 (24)	28 (22)	44 (34)	28 (25)
10 minutes to < 15 minutes	33 (24)	42 (29)	43 (32)	36 (26)	44 (32)	38 (30)	32 (25)	34 (27)	40 (32)	42 (33)	25 (19)	28 (25)
15 minutes to < 1 hour	62 (45)	58 (40)	49 (37)	50 (37)	48 (35)	44 (34)	39 (30)	37 (29)	33 (27)	23 (18)	22 (17)	27 (24)
≥ 1 hour	8 (6)	8 (6)	2 (2)	3 (2)	2 (1)	3 (2)	1 (1)	1 (1)	3 (2)	4 (3)	3 (2)	3 (3)
I do not do the exercises	4 (3)	2 (1)	1 (1)	10 (7)	9 (7)	2 (2)	14 (11)	10 (8)	6 (5)	23 (18)	20 (16)	16 (14)

Figures are *n* (%).
Analysis completed on non-imputed data as only the primary and secondary outcomes (excluding accelerometer data) were included in the imputation model.

org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com.

The imputed data set contained 514 patients and this informed the base-case analysis. Complete outcome data were available for 325 patients (64% of the total sample) and this sample was used as part of the sensitivity analysis.

Resource use

Details of NHS and total health-care resource use associated with the three trial interventions are reported in *Table 39*. Uptake of primary care and secondary care NHS services was similar between the intervention groups with the exception of knee-related surgical procedures and extra visits to NHS consultants. A higher proportion of participants in the ITE group had full total knee replacement surgery by the 18-month follow-up ($n = 10$, 7%) than either the UC or TEA groups ($n = 5$, 4%, and $n = 3$, 2%, respectively).

TABLE 39 Summary of mean resource use by treatment arm over 18 months

Resource category	UC ($n = 141$)	ITE ($n = 134$)	TEA ($n = 120$)
Primary care: GP, mean (SD) visits per patient	1.42 (2.3)	1.50 (3.2)	1.33 (2.7)
Primary care: practice nurse, mean (SD) visits per patient	0.19 (0.7)	0.37 (1.7)	0.45 (2.3)
Primary care: other professionals, mean (SD) visits per patient	0.32 (1.3)	0.28 (1.4)	0.35 (1.6)
NHS consultant, mean (SD) visits per patient	0.93 (2.3)	1.68 (3.6)	1.64 (4.2)
NHS other health-care professionals, mean (SD) visits per patient	0.12 (0.6)	0.17 (0.9)	0.04 (0.2)
Private consultant, mean (SD) visits per patient	0.50 (2.8)	0.43 (2.1)	0.11 (0.8)
Private other health-care professionals, mean (SD) visits per patient	0.45 (0.5)	0.32 (3.1)	0.00 (–)
Full knee replacement,[a] *n* (%)	3 (2)	10 (7)	5 (4)
Arthroscopy,[a] *n* (%)	6 (4)	4 (3)	4 (3)
Partial replacement,[a] *n* (%)	0.00 (–)	0.00 (–)	1 (1)
Investigations/injections,[a] *n* (%)	29 (21)	36 (27)	29 (24)
Prescribed medication,[a] *n* (%)	50 (35)	44 (33)	37 (31)
Basic analgesic	39 (22)	74 (18)	65 (40)
Moderate combination	1 (1)	2 (1)	2 (1)
NSAIDs and COX2-inhibitors	24 (14)	4 (8)	15 (19)
Strong combination opioids + opioids	19 (11)	27 (15)	25 (15)
Very strong single opioids	0 (–)	1 (1)	1 (1)
Weak combination opioids	17 (10)	19 (11)	13 (8)
Over-the-counter[a] treatments, *n* (%)	56 (40)	52 (39)	47 (39)
Basic analgesic	71 (41)	74 (42)	65 (40)
NSAIDs	2 (1)	3 (2)	1 (0.6)
Weak combination opioids	4 (2)	8 (5)	3 (2)

COX2-inhibitors, cyclo-oxygenase-inhibitors.

a The *n* (%) of participants reporting usage within the procedures, investigations, procedures, out-of-pocket and prescribed medication categories are reported instead of mean (SD) because it was not possible to distinguish multiple usage, purchases and/or prescriptions.

Adapted from Kigozi *et al.*[265] © The Author(s) 2018. Published by Oxford University Press on behalf of the British Society for Rheumatology. This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Table 40 shows the mean costs associated with the resource use during the trial. From a NHS perspective, the difference in costs between ITE and UC was £273.33 (95% CI –£62.1 to £562.6), with the ITE arm incurring higher costs to the NHS. The incremental cost between UC and TEA was £141.80 (95% CI –£135.6 to £408.1), with TEA incurring slightly higher costs. The differences were attributed to the increased number of physiotherapist treatment sessions and higher knee surgery costs in the ITE and TEA groups.

Health outcomes and quality of life

Table 41 details the EQ-5D-3L scores at baseline as well as at 3, 6, 9 and 18 months. On average, participants in all three intervention groups showed improvements at each follow-up period up to 18 months, with the exception of the ITE group at 9 months and the TEA group at 18 months. For the base-case analysis, over the 18-month period, the mean adjusted difference in QALYs between the ITE and UC groups was –0.009 (95% CI –0.068 to 0.048) and between the TEA and UC groups was –0.009 (95% CI –0.073 to 0.045). Therefore, those patients receiving UC had very slightly higher QALYs over the 18-month follow-up period than those receiving either the ITE or TEA interventions.

TABLE 40 Mean (SD) costs by treatment arm over 18 months

Items of care for costing	UC (*N* = 175)	ITE (*N* = 176)	TEA (*N* = 163)
Total cost of intervention (£)	***43.89 (20.8)***	***70.83 (32.9)***	***85.97 (41.7)***
Complete case analysis (£)	n = *141*	n = *134*	n = *120*
Primary care			
GP consultations	34.52 (55.3)	36.73 (73.8)	34.72 (71.7)
Practice nurse consultations	2.45 (8.7)	5.17 (27.1)	6.26 (33.9)
Consultations with other professionals	5.95 (34.9)	4.72 (32.4)	7.04 (35.6)
Prescriptions	7.60 (17.4)	7.64 (21.7)	4.77 (9.6)
Secondary care			
NHS consultant	67.05 (191.1)	106.33 (231.4)	103.88 (250.4)
Consultation with other NHS professional	0.00 (–)	3.28 (19.3)	1.43 (12.4)
NHS investigations and treatments	15.45 (51.6)	61.01 (513.7)	16.73 (64.3)
Knee surgery	213.79 (1194.5)	389.36 (1501.5)	259.93 (1118.6)
Private consultant	26.80 (149.2)	30.93 (166.8)	8.21 (46.1)
Consultation with other private health-care professional	0.00 (–)	13.75 (138.1)	0.00 (–)
Over-the-counter purchases	17.61 (72.3)	16.80 (44.5)	28.90 (133.3)
Base-case analysis (imputed)	n = *175*	n = *176*	n = *163*
Total NHS cost (£)	382.63 (1351.3)	656.03 (1617.1)	524.44 (1258.2)
Adjusted mean difference (95% CI)		273.33 (–62.1 to 562.6)	141.80 (–135.6 to 408.1)
Total health-care costs (£)	427.19 (1457.8)	711.11 (1683.7)	560.41 (1307.1)
Adjusted mean difference (95% CI)		283.92 (–73.43 to 591.6)	133.22 (–178.3 to 410.7)

Adapted from Kigozi *et al.*[265] © The Author(s) 2018. Published by Oxford University Press on behalf of the British Society for Rheumatology. This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com.

TABLE 41 Summary of EQ-5D-3L and QALY outcomes by treatment group over 18 months

Variable	UC	ITE	TEA
Primary imputed analysis	**n = *175***	**n = *176***	**n = *163***
EQ-5D scores, mean (SD)			
Baseline	0.636 (0.23)	0.644 (0.23)	0.629 (0.23)
3 months	0.686 (0.20)	0.708 (0.19)	0.669 (0.23)
6 months	0.690 (0.23)	0.692 (0.22)	0.692 (0.22)
9 months	0.698 (0.22)	0.665 (0.25)	0.702 (0.20)
18 months	0.700 (0.22)	0.700 (0.21)	0.682 (0.23)
QALYs, mean (SD)			
Unadjusted	1.0351 (0.27)	1.0258 (0.27)	1.0258 (0.27)
Adjusted mean difference (95% CI)[b]		–0.009 (–0.068 to 0.048)	–0.009 (–0.073 to 0.045)
Adjusted	1.0353	1.0198	1.0322
Complete case analysis	**n = *109***	**n = *113***	**n = *103***
QALYs, mean (SD)			
Unadjusted	1.0581 (0.27)	1.0478 (0.21)	1.0282 (0.29)
Adjusted mean difference (95% CI)[b]		–0.0103 (–0.083 to 0.059)	–0.0298 (–0.105 to 0.046)
Adjusted[a]	1.0614	1.0329	1.0409

a Predicted scores after controlling for baseline imbalances.
b With UC as comparators.
Adapted from Kigozi *et al.*[265] © The Author(s) 2018. Published by Oxford University Press on behalf of the British Society for Rheumatology. This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com.

Cost-effectiveness analysis

The results indicate that UC is less costly than the ITE and TEA interventions and is at least as effective (*Table 42*), and is therefore the dominant treatment. *Figures 10* and *11* show the corresponding cost-effectiveness planes for the base case. Each point results from a separate bootstrap replication, with the 5000 points scattered in all four quadrants and the majority of points located in the north-west quadrant in both cases (68% for ITE vs. UC and 45% for TEA vs. UC), showing the dominance of UC. The corresponding CEACs are represented in *Figure 12*, showing the very low probability of either the ITE or TEA physiotherapy intervention being cost-effective.

TABLE 42 Cost–utility analysis for 18 months' follow-up

Intervention	Mean costs (SD) (£)	Mean QALYs	ICER (£/QALY gained)
UC	382.63 (1351.32)	1.0353	N/A
ITE	656.03 (1617.07)	1.0322	Dominated by UC
TEA	524.44 (1617.07)	1.0198	Dominated by UC

Adapted from Kigozi *et al.*[265] © The Author(s) 2018. Published by Oxford University Press on behalf of the British Society for Rheumatology. This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

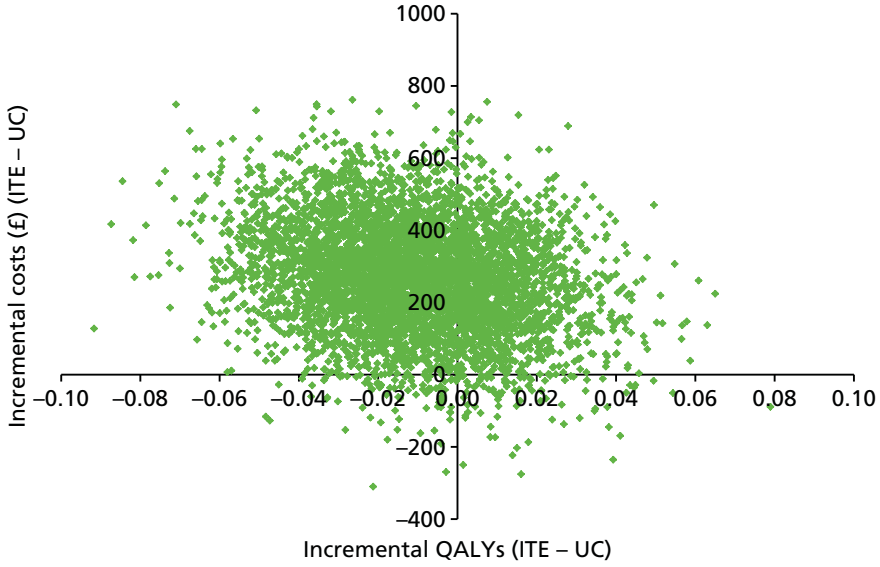

FIGURE 10 Base-case cost–utility plane comparing ITE with UC. Data were based on 5000 bootstrapped cost-effect pairs. Incremental costs and incremental QALYs were calculated as ITE minus UC. Adapted from Kigozi *et al.*[265]
© The Author(s) 2018. Published by Oxford University Press on behalf of the British Society for Rheumatology. This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com.

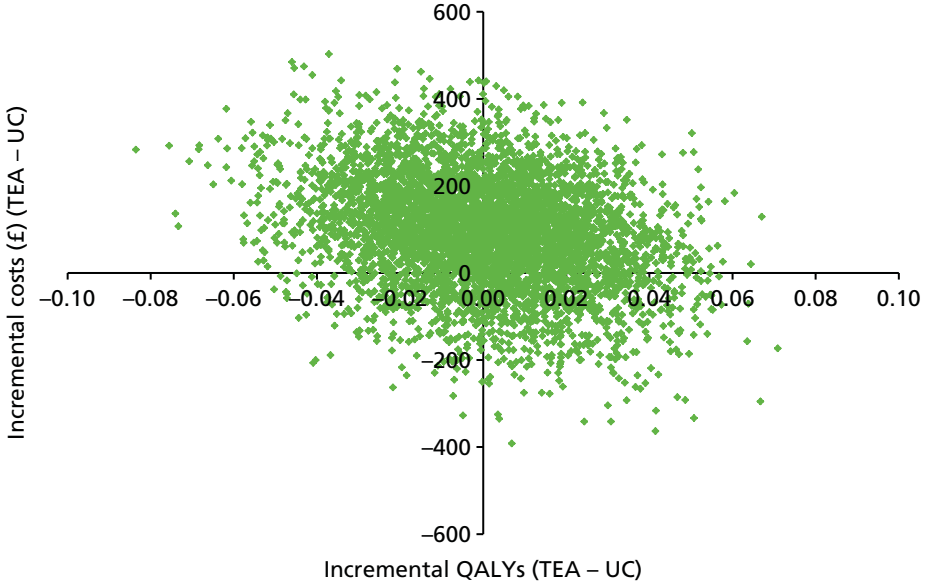

FIGURE 11 Base-case cost–utility plane comparing TEA with UC. Data were based on 5,000 bootstrapped cost–effect pairs. Incremental costs and incremental QALYs were calculated as TAE minus UC. Adapted from Kigozi *et al.*[265]
© The Author(s) 2018. Published by Oxford University Press on behalf of the British Society for Rheumatology. This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com.

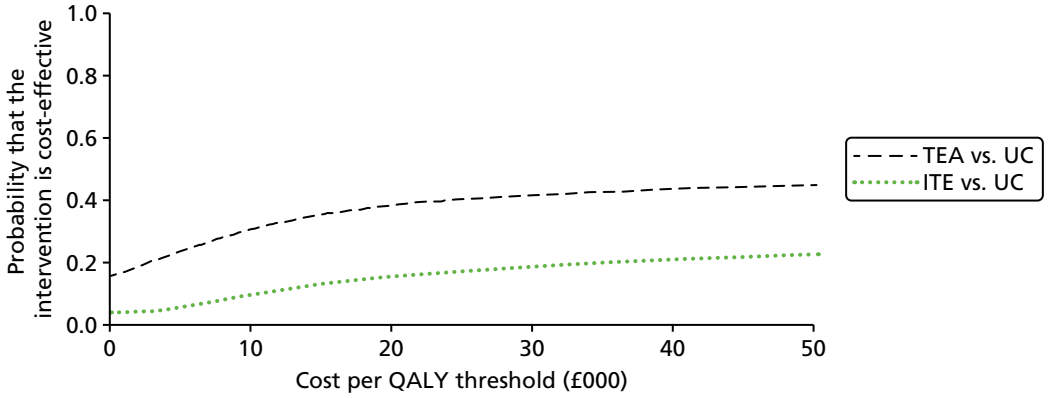

FIGURE 12 Base-case cost–utility plane comparing TEA and ITE with UC. Data were based on 5000 bootstrapped cost–effect pairs. Adapted from Kigozi *et al.*[265] © The Author(s) 2018. Published by Oxford University Press on behalf of the British Society for Rheumatology. This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com.

Sensitivity analysis

The mean difference in QALYs for the complete case analysis over the 18-month period between the ITE and UC groups was –0.0103 (95% CI –0.083 to 0.059) and –0.0298 (95% CI –0.105 to 0.046) between the TEA and UC groups, leading to similar conclusions as the base-case analysis. The incremental mean QALY estimates over 18 months were larger than the base-case analysis. The broader health-care perspective, taking into account private health care and over-the-counter expenditure by patients, showed the incremental mean health-care costs were also higher in the ITE and TEA groups in the imputed and complete case data sets (see *Table 40*). The same implications were replicated in all sensitivity analyses that varied the costing methodology, including the complete case analysis and imputation of total NHS costs (*Table 43*).

TABLE 43 Sensitivity analysis mean (SD) costs by treatment arm over 18 months

Costs	UC	ITE	TEA
Sensitivity analysis 1: complete case analysis	**n = *141***	**n = *134***	**n = *120***
Total NHS costs (£)	423.08 (1517.5)	712.01 (1762.2)	519.39 (1278.9)
Adjusted mean difference (95% CI)		288.93 (–147.5 to 656.6)	96.30 (–271.1 to 455.8)
Total health-care costs (£)	467.50 (1643.1)	773.49 (1848.1)	556.51 (1350.2)
Adjusted mean difference (95% CI)		305.90 (–178.2 to 382.4)	89.00 (–320.5 to 474.8)
Sensitivity analysis 2: multiple imputation of total NHS costs	**n = *175***	**n = *176***	**n = *163***
Total NHS costs (£)	395.36 (1324.7)	675.96 (1602.7)	463.71 (1090.08)
Adjusted mean difference (95% CI)		280.60 (–42.8 to 594.3)	68.34 (–177.7 to 322.5)
Total health-care costs (£)	437.38 (1433.4)	732.27 (1676.4)	497.78 (1150.8)
Adjusted mean difference (95% CI)		294.88 (–36.61 to 628.4)	60.39 (–206.1 to 377.9)

Adapted from Kigozi *et al.*[265] © The Author(s) 2018. Published by Oxford University Press on behalf of the British Society for Rheumatology. This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Work-related outcomes

Information regarding paid employment and productivity costs due to time off work resulting from knee pain is reported in *Table 44*. At 18 months, the percentage of people taking time off work was higher in the ITE intervention group (17.4%) and the TEA group (14.6%) than in the UC (3.3%) group. This translated to higher productivity costs in the intervention groups than in the UC group. The difference between the ITE and UC groups was statistically significant and the difference between the TEA and UC groups was not significant.

In summary, the health economic analyses showed that UC was less costly than, and at least as effective as, the other two interventions in the BEEP trial. Participants in the UC group went on to receive fewer knee-related surgeries and had fewer visits to NHS consultants and less time off work over the 18 months' follow-up.

Results from the linked qualitative study

Parts of this text have been reproduced with permission from Moore A, Holden M, Foster N, Jinks C. Keele University. 2017.[266]

In total, 30 participants were interviewed shortly following the end of their BEEP treatment. This was the first interview time point (interview 1). Of these, six (two from each arm) declined to participate in the follow-up interviews (interview 2) which occurred 12–18 months after completion of BEEP treatment (timed to occur following participants' 18-month follow-up questionnaires). Two participants withdrew from the qualitative interviews (but remained in the trial). One participant withdrew because of a bereavement while the other gave no reason. Therefore, 22 participants were interviewed at the second time point. Overall, there was a spread of participants across the three intervention groups at interview 1 (with the smallest number in the ITE group) (*Table 45*). There was an equal split according to sex across the interview sample and participants were drawn from across all age ranges (although there were fewer in the youngest age group). Overall, the interview sample included more participants reporting improved rather than worsened pain or levels of physical functioning following BEEP interventions.

Participants' perceptions of BEEP interventions

The interviews revealed patient perceptions of what happened during BEEP treatment sessions. Physiotherapy-led exercises were described across all three arms of the trial and these data shed light on how core intervention constructs of supervision, progression and individualisation were experienced. Different levels (basic and higher) emerged from the data (*Figure 13*).

TABLE 44 Description of work-related outcomes for participants in paid employment, by treatment group

Work outcome	UC	ITE	TEA
Working in paid employment at 18-month follow-up	61/141 (43%)	46/134 (34%)	41/120 (34%)
Reported time off work due to knee problem at 18 months	2 (3.3%)	8 (17.4%)	6 (14.6%)
Mean (SD) time off work due to knee problem over 18 months (days)	1.79 (13.5)	13.55 (33.5)	6.95 (24.2)
Mean (SD) cost (£) of knee-related work absence	127.84 (963.3)	1313.61 (3368.2)	691 (2382.1)

Values are *n* (%) unless stated otherwise. The estimation of indirect costs focused on the subsample of respondents in paid employment at 18 months (148/514).

Adapted from Kigozi *et al.*[265] © The Author(s) 2018. Published by Oxford University Press on behalf of the British Society for Rheumatology. This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com.

TABLE 45 Summary of characteristics of interviewees

Sampling criteria	Subcategory
Age (years) (at time of first interview)	45–55 ($n = 5$)
	55–64 ($n = 9$)
	65–74 ($n = 8$)
	≥ 75 ($n = 8$)
Sex	Female ($n = 15$)
	Male ($n = 15$)
Treatment group	UC ($n = 10$)
	ITE ($n = 8$)
	TEA ($n = 12$)
WOMAC pain[a]	Improved[b] ($n = 18$)
	Worsened[c] or no change ($n = 11$)
	Missing ($n = 1$)
WOMAC physical function[a]	Improved[b] ($n = 19$)
	Worsened[c] ($n = 11$)

a WOMAC Pain and Physical Function Scores were taken from self-completed questionnaires returned at 3 months follow-up.
b Improved = any reduction in pain or functional disability subscale scores on 3 month questionnaires since baseline.
c Worsened = any increase in pain or functional disability subscale score on 3 month questionnaires.

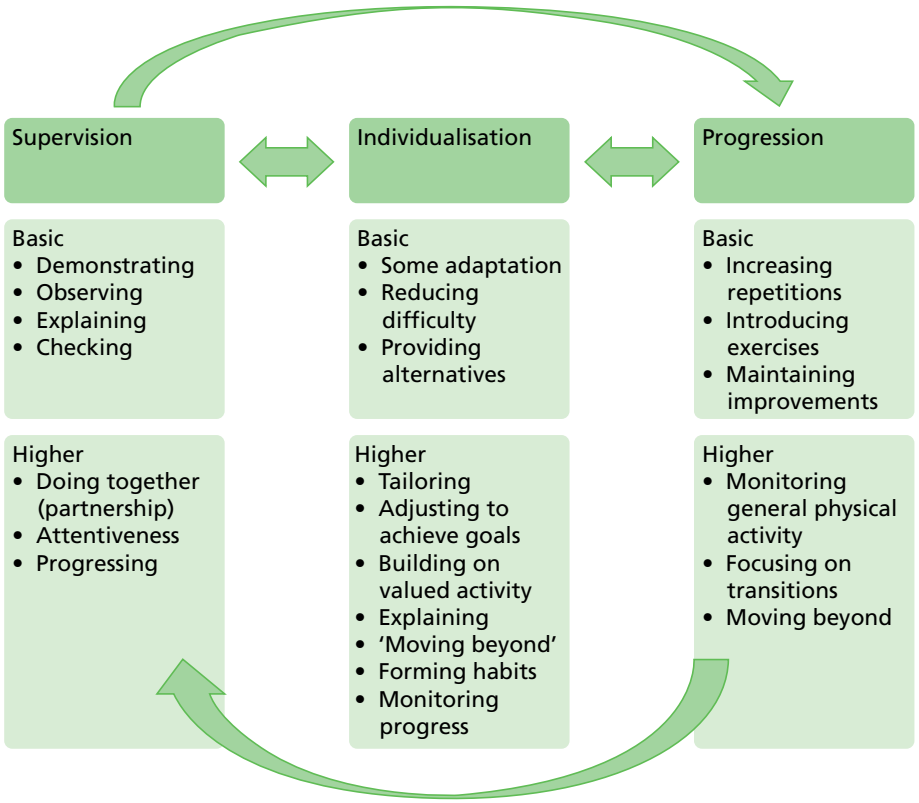

FIGURE 13 Evidence of intervention constructs described in qualitative interviews in all three intervention arms of the BEEP trial.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Participants in all arms described how their treating physiotherapist demonstrated the exercises, watched them, explained the exercises and assessed and monitored for any symptom response to the exercise programme (basic supervision). Higher levels of supervision were also evident across all three arms. However, in the ITE arm, a notable difference was the sense of partnership with the physiotherapist that participants described. This was highlighted by descriptions such as 'we practised them', 'we did them', 'we went through them together' and 'he literally went through them and I did them all with him.' In the other two arms, although participants recalled physiotherapists demonstrating the recommended exercises, often a more sequential approach occurred with the patient following the physiotherapist's demonstration.

A basic level of exercise progression included increasing the number of repetitions of exercises, introducing new exercises and maintaining improvements in exercise. This level was evident from the interviews with participants from all interventions arms of the trial. Higher levels were described in examples that participants in the ITE and TEA intervention arms gave about moving beyond specific lower limb BEEP exercises to more general physical activities. For example:

> *We went onto weight-bearing exercises then, so I'm doing sort of step ups and squats, and an exercise bike as well, and I'm using an exercise bike. So you know 12 months ago, I wouldn't have been able to try and pedal so, you know, it's – it's very good really.*
>
> *6878, ITE: interview 1*

> *I know she felt quite certain that she could get me walking around the block.*
>
> *4876, TEA: interview 1*

The interviews revealed evidence of a basic level of exercise individualisation in all three intervention arms, characterised by adaptation of exercises based on the patient's ability to perform them (at home or at the treatment centre), reducing difficulty and providing alternatives after a period of testing out at home. There was no evidence from the interviews that individualisation in the UC arm extended beyond this basic level. A higher level of individualisation was revealed in the narratives of participants in both the ITE and TEA arms and this was reflected in participants perceiving that their physiotherapist adjusted their exercise programme to suit their individual needs. For example:

> *She was adaptable in her approach, she, erm, said, 'OK, don't do that set of exercises. This is another that's not exactly dead on for what you want, but it won't affect your hernia issue'.*
>
> *7880, TEA: interview 2*

Adjusting exercises to achieve goals, providing explanations for those adjustments and exercise replacements, monitoring progress and moving beyond specific lower limb BEEP exercises were all other characteristics of a higher level of individualisation.

> *He gives you very specific exercises to deal with the problem you've been referred with. And then he moved on to thinking about the other things you did, like your cycling which you'd done in the past.*
>
> *3657, TEA: interview 1*

There were clear links between the constructs of supervision, progression and individualisation. For example, this participant talked about being observed by the physiotherapist (basic supervision) and moving on to do new exercises (basic progression):

> *And [physiotherapist] will be watching me doing all these exercises, and she say 'I know you can do that one so we'll cross that off, we'll go on to the next one'. [. . .] Oh she would watch me close, I can tell you [laughs].*
>
> *6153 UC: interview 1*

Another participant described how, after progressing well through the BEEP exercises, he was encouraged to participate in a general physical activity that was valued by him:

> *That's where he started initially to suggest doing something else and he's probably picked up from me that I love cycling.*
>
> *3657, TEA: interview 1*

The BEEP interventions included techniques and tools to help increase general physical activity. The ITE and TEA interventions were to include SMART goal-setting, corrective feedback and reinforcement, and the TEA intervention in particular included a focus on general physical activity behaviours including techniques such as behavioural contracting with the physiotherapists' toolkit. However, from the interviews, there was little evidence of participants' recalling SMART goal-setting, no evidence of a behavioural contract being used specifically, little evidence of active physiotherapy support to access local physical activity opportunities and facilities, and no evidence of the use of exercise reminders. Although the interviews were based on a small sample (30 participants, only 8 in the ITE and 12 in the TEA interventions) and the data are subject to recall problems, the results support the quantitative findings from the physiotherapy CRFs by further suggesting that the physiotherapists in the BEEP trial did not routinely use all of the tools offered to them as part of the BEEP interventions.

Facilitators of the uptake and maintenance of exercise and general physical activity

There were a range of factors that helped participants to engage in the recommended exercises or participate in general physical activity (*Table 46*). Patient attributes acting as facilitators included self-motivation, being 'naturally active', making a conscious effort, being confident with exercise and having an active identity, for example 'not the sort to sit about'. Physiotherapist attributes included reassurance, confidence building and positivity, being mature and confident in manner, being respectful, listening and giving realistic advice. Being encouraging and enthusiastic and offering explanations also aided participants to engage with their prescribed exercises or general physical activity.

A core construct of therapeutic alliance emerged, defined here as the working relationship or collaborative bond between patient and therapist.[237] A therapeutic alliance was generated in treatment sessions with participants from all arms of the trial. The narratives of participants in the ITE and TEA intervention arms contained richer accounts of the components of this working relationship and what the impact of this was. Key features of the therapeutic alliance within each of the intervention groups in the BEEP trial are outlined in *Table 47*.

The therapeutic alliance had strong links to changing knowledge, visual motivations, and physiotherapist and patient attributes, and was aided through supervision, progression and individualisation. Although physiotherapists' attributes and therapeutic alliance influenced participants directly during treatment sessions, they provided a foundation that facilitated a change in knowledge about the role of exercise, the effects of which continued into the longer term. Those in the usual physiotherapy group spoke of an increased awareness of the benefits of exercise, the need to maintain strength and stability in their leg(s) and using painkillers specifically in order to facilitate engagement in exercise. In the ITE group, participants demonstrated awareness that progress could take time and they reported having gained a better understanding of the importance of maintaining exercise and building strength to support their knee, as well as exercising even when they have pain. In the TEA group, participants spoke about an increased understanding of the purpose and benefit of the exercises and whether or not to exercise when they felt pain during exercise:

> *One thing I've learnt from the study is that the more exercise I do with it I think the better it is going to be.*
>
> *1736 TEA: interview 2*

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 46 Factors facilitating uptake and maintenance of exercise and general physical activity

Activity category	UC		ITE		TEA	
	Post intervention	Follow-up	Post intervention	Follow-up	Post intervention	Follow-up
BEEP exercises	Physiotherapist; patient (enjoyment, feeling benefit, increased self-efficacy); time and place	Patient (feeling benefit, change in knowledge, prevention motivations)	Physiotherapist; patient (change in knowledge, feeling benefit, prevention motivations, maintain valued activity); time and place; visual motivation; supervision; social support	Time and place (fit into work)	Physiotherapist; patient (change in knowledge, enjoyment, prevention motivations); time and place; visual motivation; tailoring; progression; supervision; goal–setting; regular contact	Patient (feeling benefit); moved on
General physical activity	Patient (self-motivated, change in knowledge, prevention motivation, weight loss)		Physiotherapist; patient (change in knowledge, maintain valued activity); time and place	Patient (naturally active, retained knowledge)	Physiotherapist; patient (change in knowledge, enjoyment, prevention motivations, progression with valued activity); time and place; visual motivation; support from others	Patient (change in knowledge, naturally active); regular contact monitoring by others; support from others; epiphany

TABLE 47 Features of the therapeutic alliance in BEEP trial interactions

Features of the therapeutic alliance		
UC	**ITE**	**TEA**
Equity of labour (not letting physiotherapist down, doing my bit) (7058)	Regularity aided recall and enhanced discussion (1481)	Physiotherapist inspired participant to join gym (3657)
Sense of partnership (1135)	Continuity – ongoing dialogue with same physiotherapist enabled openness, trust, rapport (6517, 6878, 1481)	Physiotherapist was respectful of whole person (3)
Negotiating exercise frequency with physiotherapist (6153)	Physiotherapist offered reassurance, 'so felt could keep going', comforting, eased concerns (6517, 6878, 1481)	Offered reassurance about pain (7880, 6436)
	Generated connection with patient, 'talked about herself', 'got to know me', 'interested in me', encouraged questions (61, 26)	Considerate of patient's issues (7880)
	Enjoyable interaction, felt at ease, relaxed, did not rush, (61, 26)	Physiotherapist was focused and attentive (7880)
	Felt valued as person (26)	Listening (7880, 6436)
	Appreciative of physiotherapy help (61)	Sense of responsibility to physiotherapist (3657)
	Reciprocity (1481)	Building an alliance, continuing relationship with a physiotherapist who could be trusted (7880)
	Offered explanations, good at explaining, full explanations (26, 6517, 1481)	
	Attentive to me – gave realistic advice, noticed weight loss, get balance right for me (26, 1481)	
	Social not medical approach	
	Support (1481)	
	Encouragement, motivating 'got pushed in a pleasant way', make it happen, external to family (1481)	
	Challenge pre-existing ideas, new thoughts, new focus (1481)	

Numbers in brackets are codes for individual interviews.

Over the longer term (by the time of the second interviews 12–18 months later), patient attributes became more important to maintaining exercise or physical activity. These included, for example, feeling and observing the benefit of exercise and motivations to avoid future surgery or reliance on medication, and were thus linked to participants' notions about how they might prevent further decline in their knee problem.

Barriers to uptake and maintenance of BEEP exercises

A range of barriers emerged from the qualitative data (*Table 48*). Patient attributes acting as barriers included lack of enjoyment of exercises, lack of will power, beliefs about limited effectiveness of exercise and low health expectations related to ageing. Participants from all three BEEP intervention arms also reported pain as a barrier to exercise at their second interviews. In all three arms, patients found that the time taken to do the exercises and the place in which exercises had to be completed adversely affected adherence. Some commented on how they found it difficult to find the time to fit the exercises into their daily life either at home or at work, others found that finding the right place to do the exercises was a barrier as they were either distracted at home or felt that they 'looked silly' doing them at work.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 48 Barriers to the uptake and maintenance of exercise and general physical activity

	UC		ITE		TEA	
Type of activity	**Post intervention**	**Follow-up**	**Post intervention**	**Follow-up**	**Post intervention**	**Follow-up**
BEEP exercises	Pain	Pain free	Patient attributes	Pain	Pain	Pain free
	Time and place	Patient beliefs	Lack of enjoyment	Drifting back	Time and place	
	Patient beliefs	Replacement	Substitution for other valued activities	Lack of monitoring	Ineffective delivery of intervention	
	No change in Knowledge		Time and place	Substitution	Lack of explanations	
	Severity of exercises			Conflicting advice from HCPs	No change in knowledge	
	Lack of supervision			Reduced self-efficacy (fear of falling)	Patient attributes	
	Physiotherapist Attributes			Time and place	No motivation when no longer seeing physiotherapist	
	Physical status				Access: proximity to clinic	
					Patient beliefs	
					Comorbidity	
					Physical status	
					Lost motivation as could not attend	
General physical activity			Time	Pain	Patient attributes	Pain
			Substituted for BEEP		Self-identity ('not natural athlete')	Weather
			Lack of enjoyment ('gym is boring')		Fear of unknown (self-efficacy)	Financial
			Lack of will power		Already active	Anxiety
			Weather		Lack of motivation	Environment: local facilities not in working order or inaccessible
					Weather	
					Comorbidity	

HCP, health-care professional.

In the UC participants, patient beliefs about the exercises were a barrier as some participants felt that their physical status was so severe that the exercises were simply ineffective. Others had received knee replacement surgery by the time of their interview and were unable to complete the exercises for a time as a result, while some experienced an increase in the amount of pain they felt during the exercises and so stopped doing them. Although there was evidence of a change in knowledge and understanding of the

purpose and nature of physiotherapy exercises in other treatment arms, a lack of change in knowledge in participants in the UC arm was perceived as a barrier – some patients perceived that the exercises were 'too easy' and, therefore, must be ineffective. One patient found the exercises 'too severe' for her, causing more pain. Therefore, the physical status of the patient as well as the patient's (mis)understandings about the nature of the exercises they were prescribed by the physiotherapist (either too easy and thus ineffective or too difficult and causing more pain) appear to be important barriers. Overall, pain status acted as a key barrier to exercise or physical activity in participants across all three arms of the trial.

Suggestions for improvements to the BEEP interventions

The main suggestion that interviewees made to improve the BEEP interventions was for more regular physiotherapy reviews or monitoring (either at 6, 12 or 24 months). This was suggested by interviewees from all three treatment arms. Regular reviews were felt to be important in order for ongoing assessment of physical status and exercise activity, to enable changes in exercises to be made appropriately and to enable physiotherapists to remind patients of why they were exercising and why it was important to continue and persist with exercise and physical activity. Participants in the UC and ITE groups felt that regular reviews with physiotherapists rather than GPs or consultants were more appropriate as they provided the right kind of advice and support and were 'easier to talk to'. Other suggestions included more treatment sessions, treatment spread over a longer time period and different modes of delivery (e.g. e-mail or telephone sessions and reminders, physiotherapists making visits to participant's homes and community-based activities).

Discussion

Summary of trial aim and key results

Any exercise programme that is carried out regularly and monitored by health-care professionals will improve pain and function in the short term.[33] We aimed to improve both the short- and long-term pain and physical function outcomes from exercise in older adults with knee pain by changing the key characteristics of physiotherapy-led exercise programmes prescribed. UC was based on up to four treatment sessions over 12 weeks and was a relatively standardised approach to lower limb exercise prescription with limited opportunity for ongoing supervision, progression or individualisation. In the ITE group, physiotherapists offered more treatment sessions (between six and eight sessions over 12 weeks) to deliver a more highly individualised, supervised and progressed lower limb exercise programme for participants, using clear monitoring tools such as exercise diaries. In the TEA intervention, the treatment period was extended from 12 weeks to 6 months and the number of treatment contacts increased to between 8 and 10, so that physiotherapists had the time to actively support participants to make the transition from a lower limb exercise programme to engage in, and maintain, general physical activities that interest them and are available locally. Physiotherapists providing the TEA intervention had an adherence-enhancing toolkit of techniques with which to individualise the way in which they supported participants to increase their general physical activity levels, including physical activity diaries, pedometers, SMART goal-setting, behavioural contracts and set-back plans. We hypothesised that the ITE and TEA interventions would lead to greater improvements in pain and functional outcomes than UC and that these improvements would be sustained for longer.

The BEEP trial results showed no statistically significant differences in the change in pain and function between treatment groups. All three groups showed improvements in pain and function and the size of those improvements in pain and function are in line with, or slightly greater than, those reported in previous meta-analyses,[33] and with the most recent Cochrane review, which reported an effect size of 0.4 for short-term pain and 0.37 for short-term function.[120] Furthermore, approximately half of all participants in the treatment groups were classified as treatment responders using the OARSI responder criteria but, again, there were no differences between the groups. Several sensitivity analyses did not change the results, although the per-protocol analysis suggested a non-significant trend towards superiority in the TEA compared with UC comparison. We showed that the moderate benefits in pain and function gained

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

during the treatment phase (3 months in the UC group and ITE intervention group and 6 months in the TEA intervention group) were maintained at longer-term follow-up time points (9 and 18 months). This was unexpected in the UC group, given the results of previous trials which have shown short-term effects that are lost over the longer term.[34]

There is no consensus on optimal methods with which to measure exercise adherence, and we used several different approaches in the trial: self-reported adherence to the exercise programme prescribed by the physiotherapist, self-reported frequency and duration of exercise, self-reported use of physical activity facilities in the last 7 days, attendance at physiotherapy treatment sessions, and measures of general physical activity (self-reported using the PASE for all participants and objective physical activity data collection with a subsample of 89 participants). Self-reported adherence was high during the treatment phase in all three groups (≥ 75% agreeing or strongly agreeing that they had completed their exercises as often as they had been advised) and remained high for longer in the TEA group than the other two groups. In addition, participants in the TEA intervention arm reported exercising more frequently than those in the ITE or UC arms. Self-reported use of local physical activity facilities increased across all treatment groups and remained higher than baseline levels, even at the 18-month follow-up. Other than a trend towards superiority in the TEA group at 9 months, the changes were not different between the groups. PASE scores increased slightly in all three treatment groups during the treatment phase, but the changes were not sufficient to make a difference to the proportion of participants meeting physical activity guidelines. By 18 months, self-reported physical activity levels and objective physical activity levels (in a subsample of participants) had dropped back to baseline levels or lower, highlighting that without ongoing monitoring or supervision from the health professional, older adults with knee pain do not sustain exercise and physical activity behaviours.

Treatment credibility was high in all three groups and remained so even at the 18-month follow-up. Confidence to exercise (or exercise self-efficacy) and outcome expectations about exercise were reasonably high at baseline and we observed small improvements in these scores in all treatment groups. In general, participants in all three groups appeared to benefit from the BEEP trial interventions over time, with all groups reporting that they felt they had greater control over their knee problem and were less concerned about their knee problem than at baseline.

Given the lack of differences in clinical outcomes, the cost–utility analysis is key to informing decision-makers about the most appropriate treatment choice. Based on outcomes over 18 months' follow-up, the results show clearly that UC is the most cost-effective treatment compared with either of the other interventions for older adults with knee pain. Both the ITE and TEA interventions were more costly than UC and resulted in slightly lower QALYs, meaning that both were dominated by UC. UC remained the dominant cost-effective treatment in sensitivity analyses. The higher costs of the ITE and TEA interventions were associated with the increased physiotherapy sessions and slightly more secondary care knee procedures (total knee replacements) within the ITE group. Time off work was also higher in the ITE group and this may be linked to the greater number of knee replacements in this group over the 18-month period of follow-up. Further data on proportions of patients proceeding to secondary care consultations and knee surgery are currently being collected at the 3-year follow-up.

Potential reasons for trial results

One potential reason for the trial results is a lack of sufficient difference between the three interventions. All participants received a thorough assessment, advice and education based on best evidence, and all were prescribed a physiotherapy-led exercise programme that was at least in line with clinical guidelines. The differences in terms of numbers and content of intervention sessions may not have been sufficient to lead to differences in patient outcomes. Fransen and McConnell[33] showed that although treatment courses of < 12 supervised sessions were effective, those with ≥ 12 provide greater effect sizes. Our trial was designed and delivered with the UK NHS and, thus, the decision about the number of treatment contacts was influenced by what physiotherapists and their managers perceived would be deliverable within the NHS, given that current practice is to offer these patients, on average, up to four treatments. We protocolised between 8 and 10 treatment contacts (this could include telephone contacts as well as to face-to-face

contacts) in the TEA interventions and, although there was evidence that physiotherapists and patients used the telephone contact option, patients, on average, received only seven treatment contacts. We do not know from this trial whether or not offering a course of ≥ 12 treatment sessions would have led to better outcomes. Although that is possible, we doubt it would be implemented within the UK NHS because the physiotherapy resource to deliver treatment on this scale is unlikely to be available. The qualitative data highlighted that, at least for those who were interviewed, participants experienced different levels of supervision, individualisation and progression of their exercise programme in ways that matched the trial protocols for each treatment arm and provided evidence that these components of the trial interventions were delivered, at least to some extent, as intended. Analysis of the physiotherapy CRFs highlighted that the number of treatment sessions differed between the treatment groups (three, on average, in UC, five, on average, in ITE and six, on average, in TEA intervention arms) and this would therefore have provided more opportunity for physiotherapists to provide a greater level of supervision, progression and individualisation of exercise for participants. In addition, the CRF data showed the high fidelity of the interventions to delivery of exercise programmes, with low use of other additional treatment approaches (mostly ice and manual therapy which is in line with our previous national survey of practice by Holden *et al.*[199]). The CRFs for the TEA intervention arm showed that only some strategies to enhance adherence and general physical activity were frequently used and these were mostly the 'simpler' educational tools (PhysioTools exercise sheets and written education materials) and behavioural tools (physical activity diaries and pedometers). The per-protocol analysis showed a trend towards superiority of the TEA intervention compared with UC in those participants who received the intervention contacts planned. Therefore, it is possible that, to some extent, the trial results are explained by a lack of intervention fidelity. The interview and CRF analyses showed that some aspects, particularly of the TEA intervention protocol (e.g. reminder postcards, clear signposting to local physical activity facilities and behavioural contracting), were infrequently used. This may indicate that the toolkit was too multifaceted and complex to use and this may have reduced the effectiveness of the intervention.

In addition, all patients in the BEEP trial were managed with physiotherapy-led advice and exercise, and received high-quality care from physiotherapists who were confident in managing knee OA and who received specific training in the delivery of the BEEP trial interventions. The analysis of physiotherapist questionnaires over time in the BEEP trial showed that even within physiotherapists who participated in the UC training programme, the following was found: use of additional interventions (such as electrotherapy and manual therapy) alongside the exercise programme reduced, there was a greater focus on general physical activity, supervision of exercise increased and the level of agreement with the attitude statements about the role of exercise for knee pain increased. This might help explain why the BEEP trial showed sustained improvements in participants' pain and function scores even at the 18-month follow-up in the UC arm.

The qualitative analysis highlighted a central role of the therapeutic alliance between patient and physiotherapist for all three treatment groups and this may help to explain the results of the BEEP trial. The interviews showed how changes in participants' knowledge fostered during interactions in treatment sessions could have an impact on an individual's exercise and physical activity > 12 months later. Therapeutic alliance was evident from the interviews in all arms of this trial and could have diluted the effect of the additional components in the ITE and TEA interventions. The influence of the therapist–patient relationship on treatment outcomes in musculoskeletal conditions was the subject of a systematic review by Hall *et al.*[267] The authors reported positive associations between therapeutic alliance and global perceived effect of treatment, as well as changes in pain, physical function, patient satisfaction with treatment, depression and general health status.[267] Our findings are supportive of the authors' recommendation for further investigation of the factors that underpin this therapeutic alliance. The interview results of our study also highlighted other factors that facilitated or inhibited exercise and physical activity similarly across all three treatment arms, and which may help explain the trial results. In line with previous research,[210,268] we found that a wide range of intrinsic (e.g. personal factors) and extrinsic (e.g. environmental) barriers and facilitators influenced exercise behaviour in this patient group. However, these factors shifted over time from an initial mix of facilitating factors that were both intrinsic and extrinsic at the end of the treatment phase to primarily patient-related (intrinsic)

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

facilitators over the longer term. All three interventions appeared to facilitate those who saw themselves as 'naturally active' to maintain levels of physical activity despite their knee pain.

A further potential contributor to the trial results is the difference in clinical experience in the physiotherapists delivering the three intervention arms. Those delivering UC who needed to attend only 1 day of BEEP training were the most clinically experienced therapists on average in the trial, as shown in *Table 22*, which shows that physiotherapists who delivered UC had a greater mean number of years in clinical practice (18.1 years) compared with physiotherapists who delivered ITE and TEA (13.8 and 8.9 years, respectively). Whether or not this may partially explain the trial results is unclear, but adjusting the analysis for the length of time of clinical experience of the physiotherapist did not change the overall conclusions. The questionnaires from physiotherapists showed greater changes in attitudes, beliefs and intended behaviours that were more likely to be sustained in the long term in those therapists trained to deliver the ITE and TEA interventions.

Finally, although the proportion of treatment responders in all three treatment arms is encouraging (at approximately ≥ 50% at both short- and long-term follow-up), this still means that about half of all adults randomised to the trial did not respond well to exercise, in terms of pain and function outcomes and overall global impression of change. Future research that helps identify more clearly who responds and who does not respond to exercise would be helpful in targeting treatment and resources.

We cannot explain why slightly more participants in the ITE intervention arm went on to receive secondary care specialist consultations and knee replacements (10 in the ITE group vs. 5 and 3 in UC and TEA groups, respectively). It may be that greater progression of the lower limb exercise over more treatment sessions led to a slight increase in the likelihood of participants and therapists discussing the merits of surgery for the individual, or that more participants struggled with the lower limb exercise prescribed for them and they sought other health-care advice as a result. We would have expected that the ITE intervention might lead to fewer secondary care consultations and fewer knee replacement surgeries, given the increase in muscle strength, movement and stability of the lower limb due to the prescribed exercise programme.

Comparison with other trials and systematic reviews

Recent systematic reviews clearly show that exercise is effective for knee OA compared with no exercise controls,[122] that exercise should be supervised[192] and that the number of sessions is important in terms of achieving greater changes in pain and function.[33,192] There is conflicting evidence from two recent systematic reviews and meta-analyses about whether combining exercise types (such as strengthening, flexibility and aerobic exercise)[122] or focusing on single exercise types are more effective for knee OA.[192] Pinto *et al.*[269] focused their review of non-pharmacological and non-surgical interventions on cost-effectiveness and showed that there is only limited evidence about cost-effectiveness of exercise and more high-quality economic evaluations are needed.

No previous trials have specifically compared physiotherapy-led exercise approaches for older adults with knee pain/OA in which key components of individualisation, supervision and progression or a focus on exercise and physical activity adherence have been included. Other large trials in knee OA have found that prolonged exercise interventions (involving up to 12 sessions of individualised, physiotherapy-led exercise regimens) provide short-term (6 months) and smaller long-term benefits (12 and 30 months) and are cost-effective compared with usual GP care.[123,124,270] We have previously shown that physiotherapy-led exercise is superior to usual GP-led care supplemented with an advice leaflet.[34] In that trial, an average of four treatment sessions was provided and this resulted in 37% of participants being classified as treatment responders at 6 months. We showed that this increased to 43% when the exercise intervention was provided over six treatment sessions and more attention was given to selling exercise to patients.[203] By long-term follow-up, these proportions had fallen to 36% and 48%, respectively. Our estimates from the BEEP trial were better at 6 months (≥ 50% were treatment responders) with similar proportions at 18 months' follow-up. Our trial results are similar to those of a recent Australian trial also comparing different exercise approaches (neuromuscular versus quadriceps muscle strengthening) for patients with

medial knee joint OA and malalignment.[271] Their physiotherapists provided 14 supervised sessions and a home programme. They showed no statistically significant between-group differences in knee adduction moment, pain or function but, as is the case in the BEEP trial, all treatment groups showed similar statistically significant within-group reductions in pain and physical dysfunction. There are some further RCTs currently in progress comparing different exercise approaches in patients with knee OA, for example Øiestad *et al.*[272] is comparing 14 weeks of strengthening exercise, cycling and control and Messier *et al.*[273] is comparing high-intensity strength training [75–90% 1 repetition maximum (1RM)], low-intensity strength training (30–40% 1RM) and healthy living education in START (Strength Training for Arthritis Trial).

Broader literature includes systematic reviews of the effectiveness of exercise referral schemes, which show small effects on increasing physical activity in sedentary adults, and calls for further research to test ways to improve adherence.[31] In addition, a UK-based large randomised trial comparing leisure centre-based exercise programmes, instructor-led walking programmes and advice alone for being physically active for inactive adults referred for exercise by their GPs showed a net increase in physical activity in all three groups with no differences between the groups.[274]

Strengths and limitations

The BEEP trial adds to the existing body of knowledge about exercise for knee OA. It is one of the largest RCTs in the field, with long-term follow-up and linked qualitative interviews and cost-effectiveness analysis. The key difference from earlier studies in the results of the BEEP trial appears to be in the maintenance of the positive effects of exercise at longer-term follow-up. At 18 months, WOMAC pain and disability scores were similar to those at 6 months and about half of all participants could be classified as treatment responders. We have also provided the first cost-effectiveness study comparing short-term outcomes with long-term outcomes from three approaches to physiotherapy-led exercise for older patients with knee pain and clearly demonstrated that the treatment of choice for the UK NHS is UC.

There are some limitations to the BEEP trial. A small proportion of patients randomised did not receive any physiotherapy treatment despite repeated attempts to engage those participants in treatment. The actual proportion (7.6%) is lower than reported 'do not attend' rates within physiotherapy services (which have previously been reported as 15%).[275] In addition, analysis of the physiotherapy CRFs showed that, on average, physiotherapists found it difficult to deliver the ITE and TEA interventions as specified in the treatment protocols, both in terms of the number of treatment sessions and some of the adherence-enhancing strategies within the TEA intervention. Despite being a key component of the TEA interventions, only about half of those participants had a clear signpost to local physical activity facilities recorded on their CRFs. A further limitation is the level of missing data on the cost outcomes, for which only 64% of participants provided complete data. However, sensitivity analysis showed the results did not differ greatly when a complete case analysis was undertaken. In addition, resource use data were requested from patients in their 18-month questionnaire, requiring recall of resource use over the previous 12 months. This length of time may have resulted in inaccurate estimates. However, this was the case for patients in all three trial arms and so it is likely any inaccuracies are balanced.

Implications for clinical practice

The BEEP trial provides no evidence that increasing individual tailoring of, and targeting adherence to, exercise and physical activity for older adults with knee pain related to OA, in the way that we did in the ITE and TEA interventions, is either more effective or less costly than UC. The results showed that all three interventions led to clinical improvements in pain and function, with no statistically significant differences between the interventions. The cost-effectiveness analysis clearly demonstrates that UC based on up to four sessions of exercise is the treatment of choice for this patient population, given the established efficacy of physiotherapy-led exercise and physical activity for patients with knee OA. These results provide guidance for clinicians in a variety of health professions who prescribe and oversee treatment and prevention of older adults with knee OA. If it is the case that some patients benefit from more intensive exercise approaches or greater support to engage in and maintain exercise over time, then further research is needed to identify those patients.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Unanswered questions and future research

Given that long-term adherence is required to maintain the benefits of exercise and physical activity, future research needs to determine how best to support older adults with knee pain to sustain engagement in exercise behaviour. Given the results of the quantitative and qualitative data from the BEEP trial, the implications for research are to further test the clinical effectiveness and cost-effectiveness of regular supervision or ongoing monitoring in particular on adherence and clinical outcomes. In addition, although the trial showed that ≥ 50% of participants responded to exercise, this meant that up to half did not. Further research that leads to better understanding and easier identification and prediction of those patients who do and do not respond to exercise would be useful to better target health-care resources and treatments. Some research is already under way proposing a preliminary clinical prediction rule to identify the few patients who may not benefit from exercise.[276] We used a range of available approaches to measure adherence in the BEEP trial and although the broad patterns were similar for most of these, the findings of some approaches appeared a little contradictory. Further research that develops better tools with which to measure exercise adherence is needed. Finally, the qualitative research highlighted the influence of the therapeutic alliance between the patient and physiotherapist in facilitating change in patient knowledge and understanding of their knee pain and the role of exercise, and in their subsequent exercise and physical activity attitudes and behaviours. Unfortunately, in this trial we did not include a measure of therapeutic alliance, although such quantitative measures are available. We did, however, collect data on perceived treatment credibility and this was high for all treatment groups. Future studies could explore whether or not therapeutic alliance is a useful predictor of outcome for this patient population.

Overall conclusions

The most recent global burden of disease study showed that hip and knee OA is the eleventh largest contributor to global disability.[1] The rising age and obesity profiles of the population mean that health-care professionals need to prepare for a large increase in demand for health-care services to treat lower limb OA. Exercise is a recommended core treatment for knee OA and the BEEP trial investigated whether or not patients' outcomes of knee-related pain and functional disability could be improved by changing the characteristics of physiotherapy-led exercise programmes to increase the individualisation, supervision and progression of lower limb exercise and to more clearly supporting a shift from lower limb exercise to general physical activity adherence over the longer term. The results show that neither the ITE nor the TEA intervention led to superior pain or functional outcomes in comparison with usual physiotherapy-led exercise and advice consisting of up to four treatment sessions over 12 weeks. Pain and disability scores improved during treatment in all three groups and there were no differences between the groups over 18 months' follow-up. Self-reported exercise adherence was high and remained higher for longer in the TEA intervention group than the other groups, but by the long-term outcome at 18 months, this had reverted to baseline levels or lower in all groups. Qualitative interview findings highlighted the benefits that BEEP trial participants gained from the physiotherapy-led exercise interventions. Although there was evidence of differences in the treatments received between the three treatment groups, there were considerable similarities across treatment groups in terms of the therapeutic alliance developed with physiotherapists and the challenges in sustaining behaviour change in the long term. The cost-effectiveness analysis demonstrated that UC is the treatment of choice for service managers and commissioners.

Chapter 5 Workstream 4: screening for distress in persons with osteoarthritis – POST

Parts of this chapter have been reproduced from Mallen *et al.*[300] © 2017 Mallen *et al.* This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

A cluster randomised controlled trial of an electronic prompt for comorbidity screening in clinical osteoarthritis consultations in primary care with nested qualitative studies

Abstract

Background: Patients with painful, disabling OA constitute a 'high-risk' group for distress, anxiety and depressive disorders. Comorbid depression and anxiety are potentially amenable to intervention and are related to adverse future course and treatment response in OA. Screening for such symptoms is recommended. Our objectives were to (1) evaluate the effectiveness of introducing GP point-of-care screening for concurrent depressive and anxiety symptoms in patients consulting for clinical OA and (2) explore GPs experiences of participating in a screening programme.

Methods/design: Pragmatic, cluster randomised, parallel trial in primary care. GP practices were randomly allocated in blocks (1 : 1) to intervention or control using a balance algorithm. Participants were patients aged ≥ 45 years consulting for OA or OA-related joint pain. In the intervention arm, GPs asked eligible patients five questions, two each on depressive and anxiety symptoms and one on pain intensity, and they were encouraged to treat OA, joint pain and depression and anxiety symptoms according to recommended clinical guidelines. In the control arm, GPs asked eligible patients one question on pain intensity and treated patients according to their 'usual care'. Patient-reported outcomes were measured at post consultation, 3, 6 and 12 months. The primary outcome was pain intensity time-averaged across all four follow-up time points. GPs experiences were evaluated by brief questionnaire and interviews.

Results: A total of 1412 eligible patients from 44 randomised practices were recruited and consented to take part (501 from 20 intervention practices and 911 from 24 control practices). For primary end-point analysis there was a significantly higher average pain score over the four follow up time points in the intervention group than the control group (mean difference 0.33, 95% CI 0.05 to 0.61; effect size 0.16, 95% CI 0.02 to 0.29). Key secondary outcomes also reflected better outcomes in the control group.

Conclusions: This trial provides evidence against a beneficial effect on patient-reported outcomes of implementing screening for anxiety and depression in primary care patients consulting for clinical OA.

Trial registration: Current Controlled Trials ISRCTN40721988.

Background

Osteoarthritis is a chronic condition affecting the joints of older people. It is one of the most common reasons for primary care consultation in the UK, with approximately 1 million people seeking treatment each year.[277] It is often associated with functional limitation and persistent pain, both of which can have an impact on mood and QoL. Despite OA being a common disorder, the ability to effectively treat and maintain an improvement in OA symptoms in primary care is limited.[278] The majority of prognostic indicators for poor outcome in OA patients are not modifiable at the point of presentation (e.g. older age, female sex and long symptom duration prior to consultation).

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Depression has been shown to occur up to four times more frequently in individuals with persistent pain conditions, such as OA, than in individuals without persistent pain,[35,37,38] and depressive symptoms are potentially modifiable.[39] Furthermore, the results of a RCT showed that treating major depression in patients with OA resulted in improvements in OA outcomes, including pain and function, and QoL.[18,279] However, the detection of depressive symptoms in primary care is poor, particularly in those patients with a chronic physical illness.[36,280]

Depressive symptoms that do not reach criteria for major depression, often termed 'subthreshold' or 'mild' depression, are still considered to be distressing and disabling, particularly for patients with a chronic disease such as OA. Screening for depression as a routine assessment endeavours to improve identification of depressive symptoms and consequently to facilitate implementation of the management of depressive symptoms. Screening for depressive symptoms during primary care consultations for diabetes and coronary heart disease patients, using two brief questions,[281] had been stipulated as part of the Quality Outcomes Framework[282] for UK primary care practice since 2006. Indeed, screening and subsequent treatment of depressive symptoms in patients with diabetes has been shown to be associated with health-care cost savings[283] and similar findings might have been expected among patients with other long-term conditions, including OA.

Literature reviews prior to the development of this trial had recommended that such screening is also carried out in high-risk groups with comorbid medical conditions,[284,285] including chronic painful conditions.[36,286] Recommendations for depressive symptom screening had been included in the NICE guidelines for holistic OA management.[15] NICE guidelines for depression in adults with a chronic physical health problem give details of how such comorbid depressive symptoms should be identified, treated and managed.[287] Ultrashort screening tools (two to three items) had been shown to be useful for identifying depressive symptoms that warrant a subsequent more detailed assessment to determine their severity.[288] Therefore, the anticipation was that identification of depressive symptoms through screening would result in a change in care to address and, when appropriate, initiate treatment for such symptoms.[289] It was this screening combined with additional interventions to provide care of comorbid depressive symptoms that appeared to be successful in reducing OA-related symptoms.[279] Screening continued to be recommended in clinical guidelines,[15] yet how beneficial such a screening and subsequent modification of care approach is for OA patients with subthreshold or mild depressive symptoms was yet to be determined.

Anxiety has also been shown to be common in primary care patients (14.6–19.5%).[290,291] However, less than half of anxious cases were estimated to be recognised by their GP.[292] It is common for anxiety to lead to the development of depression[293] or co-occur with depressive symptoms[294] in primary care patients, and to have a potentially negative impact on well-being.[295] Consequently, screening for anxiety symptoms in OA primary care patients has been recommended as the first step in improving patient outcomes and to help prevent the development of depressive disorders.[296,297] A two-item screening tool for Generalised Anxiety Disorder Assessment (GAD-2)[291] is suitable to detect anxiety symptoms in patients across a continuum of anxiety symptoms.

Therefore, we undertook a cluster RCT to investigate prompting for recommended holistic assessment of OA by screening for both anxiety and depression in an OA primary care population. A cluster RCT rather than an individual randomised trial was chosen for both scientific and practical reasons. GPs are likely to find it difficult to act differently towards patients if those patients are individually randomised to control and intervention arms during the consultation and, therefore, contamination between the two arms would be likely in such a design. The cluster RCT in this situation represents a professional cluster intervention type,[298] in that the intervention involves professional activity during consultation and, although the patient can opt out of their data being used, the intervention is still likely to have an effect on them as it involves introducing specific questions to a consultation for diagnosed OA or peripheral joint pain suspected of being associated with OA. [This is the identical set of Read codes to those used in the consultation studies of workstream 2 (see *Chapter 3*).]

The purpose of this workstream was to investigate the effect on clinical outcomes [pain intensity and a measure of the extent to which pain interferes with daily activity ('pain interference')] of introducing structured prompts to alert GPs to the possible presence of concurrent depressive and anxiety symptoms as part of a holistic assessment in patients presenting to their GP with OA, as recommended in a systematic review of literature regarding screening for depression in primary care.[299] The qualitative work aimed to explore how useful GPs consider case-finding for anxiety and depression in people with OA, and what the possible barriers to and facilitators of such case-finding might be.

Aims

The overall aim of this trial was to develop and test a systems intervention to prompt holistic OA treatment in primary care, specifically a screening tool for concurrent depressive and anxiety symptoms.

Primary objective

The primary objective was to determine whether or not screening for generalised anxiety and depressive symptoms in patients who consult their GP with OA or peripheral joint pain results in an improvement in their pain-related outcomes. The specific hypothesis was that the patients who are screened for depressive and anxiety symptoms in their GP consultation will have a larger improvement in their self-reported current pain intensity and pain interference with daily activity ratings (NRS 0–10) over the 12 months following their consultation compared with the observed improvement in these measures among patients in the control arm (who are not screened).

Consistent with our overall aim, an integrated qualitative study based within participating general practices aimed to:

1. explore GP experience and the perceived value of including questions about anxiety and depressive symptoms in the OA consultation
2. investigate how GPs engaged with the trial and in the use of the screening template during their consultations, and whether or not there were differences between the two arms of the trial.

Study protocol: methods and design

Full ethics approval was given for this study from The Black Country REC (reference number 11/WM/0093).

Study design: main trial

The Primary care Osteoarthritis Screening Trial (POST) was designed as a cluster RCT with GP practices as the unit of randomisation. The intervention component of the trial took place during a consultation for OA. Following consultation, individual patients were invited to take part in a post-consultation postal questionnaire and subsequent follow-up questionnaires at 3, 6 and 12 months following completion of the post-consultation questionnaire. A summary of the study design is illustrated in *Figure 14*. The first assessment is 'post consultation' and is considered an outcome time point as opposed to a true baseline. A real clinical effect is possible at this early stage given the nature of the intervention both in terms of GP behaviour and medication management. Thus, four outcome time points apply to this study relating to (1) post consultation, (2) 3 months, (3) 6 months and (4) 12 months outcomes. The 'summary outcome' pain score (primary end point) concerns the average score across all four time points.

Participants

GP practices A total of 44 Royal College of General Practitioners-approved 'research ready' GP practices that use the EMIS consultation system were recruited through Primary Care West Midlands North PCRN. The GP practices involved in the study included those from both urban and rural areas, different socioeconomic groups and all sizes of practice (based on patient list size and number of GP partners). It was anticipated that the template would be active in each GP practice for approximately 3 months. GP practices were recruited over a 1-year period and so the template was not active in all practices at any one time.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

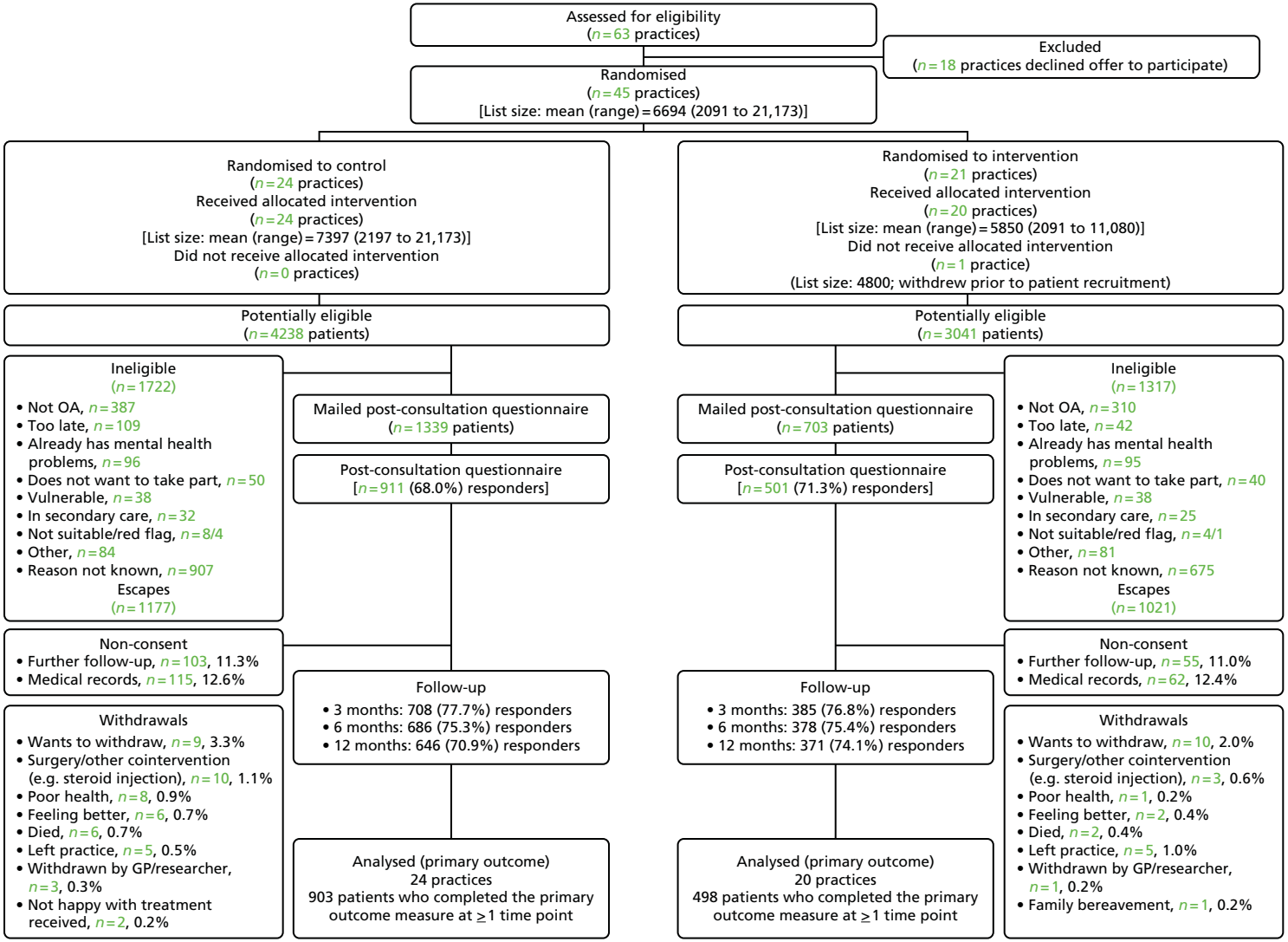

FIGURE 14 Flow chart of main study design. Reproduced from Mallen *et al.*[300] © 2017 Mallen *et al.* This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Individual patients Individual patients aged ≥ 45 years who consulted their GP with OA [defined from an identical designated set of pre-determined Read codes to those described in workstream 2 (see *Chapter 3*)] formed the sampling base for the trial. The Read code list was developed by consensus using GPs with expertise in musculoskeletal medicine. They included codes identified as being OA-specific, as well as symptom codes commonly used by GPs to identify patients with OA-related joint pain.

Inclusion criteria:

- Aged ≥ 45 years.
- Registered with the participating GP practices during the specified study period of that practice.
- Read-coded OA consultation within the specified study period (termed the 'index consultation'; may be first, new episode, or ongoing problem).
- Provided full written informed consent to study participation and to further contact at time of post-consultation questionnaire completion.

Exclusion criteria:

- Patients under active care for, or who have had a diagnosis of, depression and/or an anxiety disorder in the past 12 months.
- Vulnerable patients, including any patients on the Quality and Outcomes Framework (QOF) mental health register, or those who have a diagnosis of dementia or a terminal illness.
- Patients who reside in a nursing home.
- Red flag pathology – recent trauma associated with clinically significant injury: acute, red, hot swollen joint.
- Inflammatory arthropathy, crystal disease, spondyloarthropathy and polymyalgia rheumatica.

On behalf of the practice team, members of the informatics staff from Keele's Institute of Primary Care and Health Sciences (IPCHS) downloaded the names and addresses of eligible patients with a completed template on a weekly basis. A study pack (a letter from their GP practice introducing the study, a patient information leaflet and a post consultation questionnaire, including a consent form) was mailed to all eligible patients. The information provided to patients was the same for both arms of the trial, helping to eliminate contamination bias.[298]

Informed consent

As cluster randomisation occurred before individual participants were identified, each GP practice consented, as 'guardians' for the patients in their care, that they were willing to both enter the trial and be randomised into either arm of the trial.[301,302] Individual consent for entry into the study and for follow-up was obtained from each patient post randomisation. Individual patients were asked at the first postal contact to consent to taking part. This contact incorporated the post-consultation questionnaire, plus sections requesting consent for further contact and for their medical records to be accessed.

Intervention

The GPs in the practices in both arms were prompted to ask a standard set of questions to OA patients during the consultation; however, the content of the screening template varied between the two trial arms. *Figure 15* shows a snapshot of the screening template for both arms. The template in the intervention arm included a total of five questions (two each for depression and anxiety screening and a fifth question on pain intensity), whereas the template for the control arm included only the one question on pain intensity. Having information on pain intensity at consultation in both arms of the trial provided a further measurement to check that a balance between the intervention and control arms is achieved. Further details of the screening template are provided in *Activation of the screening template*.

Intervention arm In order to maximise 'best practice' in addition to the screening template in the intervention arm, GPs were referred to and advised to follow NICE recommendations for the holistic assessment and treatment of OA for all patients consulting with OA.[15] At a post-randomisation meeting

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

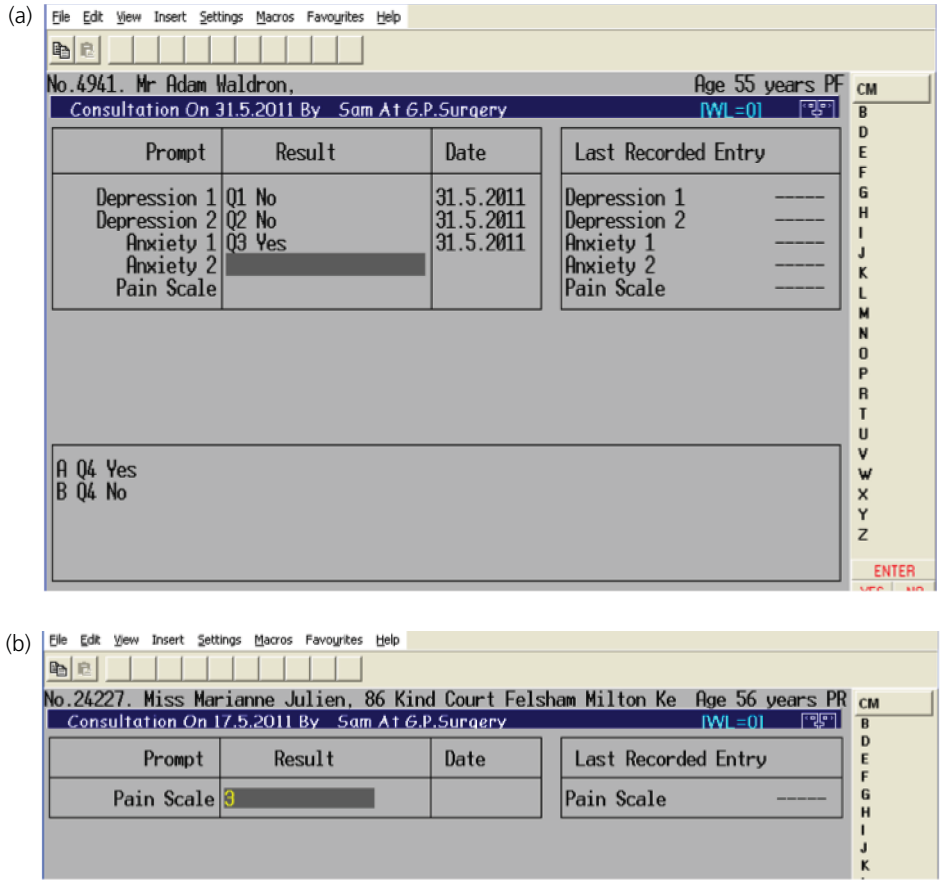

FIGURE 15 Snapshots of electronic screening template for (a) the intervention arm; and (b) the control arm. (Note: the names are fictional).

approximately 1 week prior to the template being activated in the practice, a GP research facilitator employed by the PCRN explained and discussed the intervention and study procedures with GPs and practice staff. Brief training was provided, explaining NICE-recommended evidence-based approaches to managing comorbid anxiety and depression,[287,303,304] and hard copies of the screening questions and quick reference versions of the guidelines were placed in all consulting rooms in the intervention practices. The control condition was not disclosed to intervention practices.

Control arm In order to keep the control arm as close to 'usual care' as possible, GPs were advised to follow their usual approach for responding to a patient's pain intensity rating. Asking a patient about the intensity of their pain is common practice in a primary care musculoskeletal pain consultation and should have little effect on the 'usual care' that the GP provides. No additional information on OA best practice was disseminated to control practices by the research team.

During the consultation in both arms of trial, the GP provided patients with a short information postcard introducing the study; therefore, patients were aware that they would be contacted by the research team at a later date.

Activation of the screening template In both arms of the trial, the electronic screening template was automatically activated when the GP entered a specific Read code (from the pre-determined study list) into the electronic patient record. This only occurred at the first time such a Read code was entered for a patient during the study period, ensuring that each patient was only sampled once. At this stage, GPs applied the exclusion criteria. Activation of the main template prompted the GP to ask the patient the questions on the template and to record their answers. The consultation then continued as the GP saw fit.

Outcomes

Primary outcome Pain severity for consulting site of pain (index pain) across the 12-month study period. The Chronic Pain Grade was included in each of the postal questionnaires,[229] which includes the question 'How would you rate your pain on a 0–10 scale at the present time where 0 is "no pain" and 10 is "pain as bad as it could be?" '. This question is identical to the pain intensity question included on the consultation screening template for both arms of the study.

Secondary outcomes Pain interference with daily activities was assessed on a NRS of 0–10. General health status was assessed by the SF-12.[171]

Detailed assessment of depressive and anxiety symptoms was conducted using the self-complete Patient Health Questionnaire 8 (PHQ-8) and GAD-7, respectively. The PHQ-8 is derived from the Personal Health Questionnaire 9 (PHQ-9), a validated 9-item measure of depression designed to both diagnose and assess the severity of depression.[239,240] The GAD-7 was established using both *Diagnostic and Statistical Manual of Mental Disorders*-Fourth Edition criteria for general anxiety disorder and items from previous existing anxiety scales.[239] As with the PHQ-8, the GAD-7 enquires about symptoms experienced over the past 2 weeks. The first two items on the PHQ-8 and GAD-7 comprise the Patient Health Questionnaire 2 (PHQ-2) and GAD-2, respectively, which were used to screen for depressive and anxiety symptoms in the intervention arm of this trial.

Additional items included in the questionnaire gathered information about participants' index pain symptoms, other sites of pain, pain catastrophising scale,[305] general medical health (including comorbidities and medication use) and demographic details. *Table 49* shows full questionnaire content at each time point.

Follow-up and mailing procedure

All eligible patients were sent a study pack as outlined above. Those who did not respond within 2 weeks were contacted again with a postcard reminder. At 4 weeks, those who had not responded were sent a second study pack and reminder letter. Participants who returned a post-consultation questionnaire and consented to further contact were sent a follow-up questionnaire at 3, 6 and 12 months following their index consultation. In order to maximise participation a short questionnaire, including only minimal data collection items (outlined in *Table 49*), was sent to participants who had not responded to the full postal questionnaire at 6 weeks. At 6- and 12-month follow-up points, an additional telephone call was used at 8 weeks (for those participants who provided a telephone number in their post-consultation questionnaire) to attempt to contact non-responders in order to get information on their pain intensity at that time. Three telephone call attempts were made to contact each participant.

Medical record review

A medical record review took place for participants who gave their consent (at the time of the post-consultation questionnaire) to have their records reviewed. Data from the consultation screening template was downloaded, alongside the date of consultation, general practice and index Read code information. After the 12-month follow-up period, a more detailed review was carried out to assess prescriptions, referrals for further treatment and number of consultations, and whether or not these were for OA or mood problems. This was used to determine whether there had been any previous diagnosis or treatment for depressive or anxiety symptoms. This data provided primary health-care resource use for interpretation of clinical effects observed and for the health economic analysis. These data are not presented in this report and will appear in future publications.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 49 Summary of full postal questionnaire content

Conceptual domain	Operational definition	Empirical measure	Number of items	Time
Primary outcomes				
Pain intensity	Current intensity of the index pain	NRS 0–10[a]	1	PC, 3FU, 6FU, 12FU
Secondary outcomes				
Disability	Composite characteristic pain interference with daily activities, recreation and work during the past 3 months	Number of days (one item) and NRS 0–10 (three items provide composite disability score of 0–100)[a]	4	PC, 3FU, 6FU, 12FU
Characteristic pain intensity	Composite current and past and average pain intensity during the past 3 months	NRS 0–10 (three items provide a composite characteristic pain intensity score of 0–100)[a]	3	PC, 3FU, 6FU, 12FU
Index pain location	Site of index pain complaint	Choice of anatomical site	1	PC
Nature of onset	Traumatic onset	Yes/no/unsure	1	PC
Episode duration	Time since last whole month free from this pain	Episode duration	1	PC, 3FU, 6FU, 12FU
Days in pain	Days in past 6 months with index pain	No days/1–30 days/ 31–89 days/90 + days	1	PC, 6FU, 12FU
Multiple site/chronic widespread pain	Pain in sites other than the index site over the last month. Manchester definition of widespread pain[264]; Keele definition of number of pain sites; chronicity of pain; pain in the knee or hip	Yes/no; body manikin	4	PC, 3FU, 6FU, 12FU
Pain consultation	First consultation for this pain complaint	Yes/no	1	PC
Pain improvement	Improvement in pain at index pain site	6-point Likert scale	1	3FU, 6FU, 12FU
Further pain consultation	Consulted since index consultation with the same pain site	Yes/no	1	3FU, 6FU, 12FU
Satisfaction with consultation	Satisfaction with questions asked during consultation	Yes/no/unsure; details; list of items asked	4	PC
Catastrophic coping	Catastrophic coping relating to index pain	Pain catastrophising scale	9	PC
Anxiety	Generalised anxiety symptoms in past 2 weeks	GAD-7	7	PC, 3FU, 6FU, 12FU
Depression	Depressive symptoms in past 2 weeks	PHQ-8	8	PC, 3FU, 6FU, 12FU
Discussion of mood in consultation	Did doctor ask about mood during index consultation	Yes/no/unsure	1	PC
Diagnosis of depression	Ever diagnosed by doctor as depressed	Yes/no/unsure	1	PC
Diagnosis of anxiety	Ever diagnosed by doctor with an anxiety disorder	Yes/no/unsure	1	PC

TABLE 49 Summary of full postal questionnaire content (*continued*)

Conceptual domain	Operational definition	Empirical measure	Number of items	Time
QoL measures				
General health	Mental and physical well-being	SF-12	12	PC, 3FU, 6FU 12FU
General health	Utility-based QoL	EQ-5D-5L	5	PC, 3FU, 6FU 12FU
Comorbidities	Ever been diagnosed with any from a list of possible comorbidities	Yes		PC
Fractures	Ever fractured hip/wrist/other bones	Yes/no		PC
Falls	Falls in the past 12 months	Yes/no	1	PC
Health-care costs				
Medication usage	Select from list of analgesic and psychiatric medicines that currently taking			PC
Health-care resource use	Use of non-primary care health-care resources	Yes/no and details of resource use if yes	3	PC, 6FU, 12FU
Over-the-counter medication	Over-the-counter medication expenditure	Details of over-the-counter medications		6FU, 12FU
Miscellaneous				
Age	Age at index consultation	Date of birth	1	PC, 3FU, 6FU, 12FU
Sex	Sex	Male/female	1	PC, 3FU, 6FU, 12FU
Employment status and absence from work	Employment status at time of questionnaire and days of work absence	Yes/no and details	10	PC, 6FU, 12FU
Socioeconomic status	Occupational class based on individual (1) current or (2) most recent job title	Job title – categorised as manual/non-manual	2	PC
Living arrangement	Live alone	Yes/no	1	PC
Marital status	Marital status at time of post-consultation questionnaire	Married/single/divorced/ widowed/separated/ cohabiting	1	PC
Social support	Availability of instrumental and emotional support (yes/no/no need)	Single items	2	PC, 6FU, 12FU
Obesity	BMI	Height (metres/feet, inches), weight (kg/stone, lb)	2	PC

3FU, 3-month follow-up; 6FU, 6-month follow-up; 12FU, 12-month follow-up; PC, post consultation.
a Seven items form the Chronic Pain Grade.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Sample size

In order to estimate the time-averaged difference in the pain response between the two arms, we calculated the number of subjects needed per arm as:

$$M = 2(z_\alpha + z_\beta)^2[1 + (n - 1)\rho] / (n\Delta^2), \quad (1)$$

using a significance level of α, power 1-β, Δ = smallest meaningful difference in SD units, n = number of repeated measures.[306] We took the effect size as an average of 0.2 over the four follow-up time points (post consultation, 3, 6 and 12 months). Using $\rho = 0.5$ as an estimate of the autocorrelation, $n = 4$, $\Delta = 0.2$, $\alpha = 0.05$, $\beta = 0.10$, gives 328 per arm. To adjust for the clustering of the practices, we modified this figure by the design effect for unequal cluster sizes:

$$1 + [(cv^2 + 1)m - 1] \times ICC.[49] \quad (2)$$

On interim investigation of the first two blocks of practices (13 in total), the ICCs for key outcomes at 3 months were < 0.005 (a value of 0.015 was deemed a sufficiently conservative estimate for the ICC). The average cluster size (m) was anticipated to be about 30 (incorporating approximately 70% participation rate from 43 approached per practice) and the coefficient of variation (cv) to be about 0.5. Hence, the design effect was calculated to be an inflation factor of 1.55; thus, requiring 510 for analysis per arm. Allowing for 20% dropout, the total sample size requirement was 1280 (from 44 GP practices based on an average per-practice recruitment of 30 participants).

Randomisation procedure

Allocation sequence generation GP practices were randomised in blocks using the balance algorithm for clustered randomised trials.[307] Baseline covariate data included PCT, GP practice list size and the Index of Multiple Deprivation. GP practices were recruited over the period of 1 year so that blocks would form as GP practices agreed to take part. The algorithm generates a set of allocations with the smallest imbalance statistic across the two arms.

Allocation concealment Final allocations of GP practices were randomly selected by the independent statistician on the Trial Steering Committee. This information was passed to the PCRN which was able to install the appropriate template into each practice and arrange for a GP research facilitator to meet with each practice to introduce the screening template and study procedures.

Blinding The principal investigator, chief investigator, trial statistician and members of the administration team, who input data from the study questionnaires, as well as the GPs, were blinded to cluster allocation. The Trial Steering Committee and Data Monitoring Committee were also blinded to cluster allocation unless it became absolutely necessary to reveal allocation. Individual patients were unaware as to which arm of the trial they were in.

Study administration

Two secure databases were designed specifically for use in this study. One was a mailing database in which all details downloaded from practice databases were stored. A unique study number was applied to each potential participant. The index pain site that triggered the initial consultation, noted in the post-consultation questionnaire, was entered into the mailing database so that follow-up mailings could refer to the index pain site. Following the reminder period of the mailing procedure, eligible patients who did not respond to the post consultation questionnaire were recorded as non-responders in the mailing database and not contacted again at any stage of the study. The name and address details of non-respondents were erased after a reasonable amount of time. The NHS tracing system was used to search for any non-responders to ensure a correct postal address.

The second database was used to store all data collected during the study. In this database, participants were only identified by study number. Data were entered into both databases by trained members of the administration team blinded to cluster allocation. Access to the databases was restricted to those members of the team that required access to this data. The coding schedule for the study questionnaires was used to inform database design and to facilitate data entry. The data entry of a randomly selected 1 in 10 questionnaires was checked by a member of the administration team who had not been inputting data from this study and who was blinded to cluster allocation. Throughout the entire mailing process, checks were carried out at each practice, by a member of the PCRN, to determine any patient deaths and departures and to ensure that there was no inappropriate contact of any patient. Questionnaires and study consent sheets were stored securely in separate locations to ensure participant confidentiality.

General practitioner engagement

In order to maintain GP engagement throughout the screening period, the study co-ordinator made regular contact with the practice manager to update them on the number of patients they had screened. At a minimum, a 3-week phone call, mid-point detailed update and 10- to 12-week phone call took place. If the number screened was lower than expected for that practice, then more regular communication was initiated and the GP research facilitator was asked to contact the practice and ensure that template training was sufficient.

Statistical methods

The main analyses were completed for the primary and secondary outcome measures based on the total study population (i.e. all those who completed and returned the post consultation questionnaire). For the primary outcome, current pain intensity (0–10 numerical rating scale), a linear mixed model was carried out allowing all available data to be analysed and estimates of effect calculated for all individuals providing outcome data on at least one occasion. Similarly, for all secondary outcomes, a linear or logistic mixed model (as appropriate to the scale of the outcome data) was performed on all available data for the study population in question. Analysis was performed on the basis of intention to treat and evaluation was undertaken as per cluster randomised allocation.

A sensitivity analysis of the primary outcome was performed on a complete case subset of the mailing response study population who provided consent to further contact and medical record review and provided complete outcome scores for the primary outcome at all four outcome time points (post consultation, 3, 6, 12 months) as well having a recorded pain template score (i.e. a ‘true’ baseline pain score). This allowed a more inclusive baseline adjustment set of covariates for examining the robustness of estimates of the primary outcome to possible confounding (notably in relation to potential imbalance in pre-treatment NRS pain score), and assessment of the robustness of the effect estimates of the primary outcome to different analytical models (and implicit model assumptions), specifically with respect to attrition bias.

A further sensitivity analysis of the primary outcome focused on intervention per-protocol analysis using CACE estimation of the effect of the intervention (among GPs who comply with protocol). Compliance is strictly specified as a necessary ‘yes’ response recorded in the template to either of the two depression items, ‘yes’ to either of the two anxiety items, or a ‘no’ response to both items in each – all other combinations imply that the template was not sufficiently completed to aid in any diagnostic screening of anxiety/depression.

As the data were collected from a cluster trial, coefficients were estimated using robust standard errors to allow for clustering. This involved analysis using a three-level hierarchical model to take account of clustering by practice and repeated measurements on individual participants. Estimated mean response profiles between the two groups were compared: (1) across all follow-up time points simultaneously as an aggregated summary (primary objective) and (2) across all time points distinctly (secondary objective to determine whether effect differences, if any, are consistent or are different at the different time points of assessment).

© Queen’s Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

The main analysis was adjusted for patients' age and sex, the time interval between consultation and questionnaire response, and the following practice-level variables: PCT, practice index of deprivation score, total practice size, consultation rate in the 12 months prior to randomisation for OA and related codes among patients aged ≥ 45 years, proportion of potentially eligible participants mailed, and proportion of mailed participants who responded to post-consultation questionnaire. Current pain (recorded on template at the first GP consultation) was only available for the subset population who consented to medical record review; this measure was a baseline covariate in the complete case sensitivity analyses of the primary outcome, along with duration of complaint and self-reported BMI.

Subgroup analyses focused on two evaluations for main investigation: age and severity of pain according to template completion (based on median cut-off point). Effect estimates were derived through additional modelling of the interaction term in regression analysis. Age cut-off points of 45–64 years, 65–74 years and ≥ 75 years were used and were specified a priori.

Health-care use

Primary health-care resource use (GP and practice nurse consultations/home visits) and medications were collected from a review of primary care medical records, covering the full 12-month follow-up period for study participants who consented to a review of their medical records. These data are not included in this report.

Potential bias

A number of steps have been taken in the design of this trial to minimise threats from selection and recruitment bias, including ensuring that GP practices were given the same information about the study prior to randomisation, patients in both arms of the trial were given the same information, the template was active in each GP practice for a relatively short length of time (3 months) and the screening template was automatically activated and opting out of template completion was recorded, including the reasons for opting out. The trial commenced in July 2011. A 3-month 'run-in' period in four GP practices (two in each arm) ensured that methods and key assumptions worked successfully and that the day-to-day administration of the trial ran smoothly. Data from this phase of the trial was included in the analysis.

Study design: qualitative studies

General practitioner questionnaire

At the end of the screening period, GPs were asked to complete a brief feedback questionnaire. This was a seven-item questionnaire (*Table 50*). Copies of the questionnaire were given to the practice managers to distribute at the end of their practice's study period (approximately 3 months). The questionnaires were

TABLE 50 Content of GP's experience of POST questionnaire

Item	Item description	Item response
1	Impact on consultation time	5-point Likert scale
2	Amount of additional time	7 options plus free text option
3	Impact on doctor–patient communication	5-point Likert scale plus additional 4-point Likert scale when there was an impact
4	Effect on patient management	5-point Likert scale
5	Reasons for excluding patients	Free text
6	Ease of incorporating screening questions	4-point Likert scale
7	Additional comments	Free text
Name		If happy to provide
Contact details		If happy to provide

not addressed to specific GPs because of the likelihood that there would be locums or new staff at each practice during the study period. The back page of the questionnaire asked GPs to consent to further contact by providing their contact details. Completed questionnaires were returned to the research centre in the stamped address envelope provided.

When a GP questionnaire was received at the research centre, the practice number was retained so as to identify which arm of the trial the GP had participated in and a study number was assigned to each GP. The information in the questionnaire was then entered into a database with all other data anonymised. The back page of the questionnaire with any contact details was removed from the hard copy questionnaire and stored separately in a locked filing cabinet.

General practitioner interviews

A subsample of GPs from across both arms who consented to further contact (in the questionnaire) were invited to take part in an interview to explore their perspectives and experiences of participating in the trial in more detail. Interviewers ensured that the study team remained blind to practice allocation. The focus was on investigating GP views on the detection and management of anxiety and depressive symptoms generally and in particular the value of including specific questions around these in the OA consultation. It also explored training and support needs in addition to their views on taking part in the trial at a practical level and their use of the screening template, and examined whether or not there were differences between the two arms of the trial.

Applied thematic analysis with constant comparison was used to analyse the interview data. In order to ensure reliability, a minimum of two members of the study team independently coded the data before agreeing a coding frame.[308] The questionnaire data was descriptively analysed. In addition, findings from the qualitative study were mapped against trial data, looking for differences and similarities in treatment, taking into account scores for the five questions, sociodemographic characteristics, comorbidities and repeat attending.

The interview schedule was designed to complement the questionnaire and to seek more in-depth clarification on responses in-keeping with a mixed-methods approach.[309] In addition, GPs from both arms of the trial were asked about the acceptability in general of case finding for patients with OA/joint pain. However, those in the control arm did not ask the PHQ-2 or the GAD-2 and, therefore, questions relating to the impact that these had on consultation time and doctor–patient communication were limited to GPs in the intervention arm. Both arms asked the question regarding pain intensity and questions regarding the impact that this had were included in the interview schedule for both groups.

To give as much flexibility as possible and to minimise the intrusion into practice time, telephone or in-person interviews were offered as well as group interviews, and at a time and place to suit participants. Copies of the information letter and consent form were e-mailed both at the time of the initial invitation and again just ahead of the interview. Consent was again checked at the beginning of the interview and audio-recorded. An outline of their content is given in *Table 51*.

Sampling

A convenience sampling approach was adopted, that is, one that involves drawing samples of persons who are both easily accessible and willing to participate.[310] Consequently, while being mindful of the need to sample equally from both arms, any GP who consented to further contact was invited to interview until data saturation was achieved (see *Results*).

Analysis

The questionnaire data were analysed descriptively. All audio files were transcribed and the data cleaned, that is, transcripts were checked against the audio files for accuracy and anonymity. A thematic framework

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 51 Outline of interview schedule for GPs

Domains	Intervention arm (yes/no)	Control arm (yes/no)
General patient population including prevalence of psychological problems	Yes	Yes
Views of psychological problems in joint pain/OA	Yes	Yes
Presentation/detection of psychological symptoms in such patients	Yes	Yes
Management	Yes	Yes
Communication	Yes	Yes
Training and development	Yes	Yes
Taking part in POST	Yes	Yes
Ease of asking questions	Yes	No
Differences between GAD and PHQ	Yes	No
Usefulness of screening template	Yes	No
Patients included/excluded	Yes	No

GAD, Generalised Anxiety Disorder.

approach to analysis was adopted as this allowed a focus on the areas covered by both the questionnaire and the interview topic guide while also allowing unanticipated themes to emerge.[308,311] Analysis was led by Bernadette Bartlam, with input from other coding team members (CC-G, CM, JR, DG and BN) as a means of providing intercoder reliability and generating interdisciplinary perspectives. In the initial stage of analysis, team members read a random selection of transcripts from each arm to familiarise themselves with the data and independently identify preliminary themes. A coding frame was developed as a result, which included a priori themes such as impact on the consultation, as well as any unanticipated themes.[311] Team members discussed these preliminary themes and developed a more detailed coding frame, which was applied to a small number of transcripts. The findings were then discussed further and a final coding frame agreed.[308] Data were searched for additional confirmatory or challenging evidence within and between individual interviews, across the questionnaire responses, when available, and across the two groups in a process of constant comparison.[184,312] The process of analysis was on-going from the first interview, and unanticipated responses were checked out for confirmation or contradictions in interviews with subsequent participants.[308]

Results

Participant flow

A flow chart illustrating the flow of centres and individual participants through the trial is given in *Figure 14.*

A total of 45 GP practices were randomised with an overall mean list size of 6694 patients. A total of 24 practices were randomised to the control group and 21 practices to the intervention group (one practice in the intervention arm withdrew prior to patient recruitment). A total of 7279 patients were identified as being potentially eligible for the trial; 3039 were deemed to be ineligible (reasons are shown in the flow-diagram). The GPs avoided using the template for 2198 patients, but 2042 patients were screened and posted a post-consultation questionnaire. The proportion of potentially eligible patients that were screened and posted a post-consultation questionnaire by the GP was higher in the control group ($n = 1339$, 31.6%) than in the intervention group ($n = 703$, 23.1%).

A total of 1412 (69.1%) participants responded to the post-consultation questionnaire: 911/1339 (68.0%) in the control arm and 501/703 (71.3%) in the intervention arm. The time between the date of consultation and date of returning the post-consultation questionnaire was a mean of 24 days (IQR 17–35 days; range 9–149 days) in the control arm and 22 days (IQR 16–33 days; range 3–106 days) in the intervention arm. Follow-up rates were similar in both arms, totalling 1093 (77.4%) at 3 months, 1064 (75.4%) at 6 months and 1017 (72.0%) at 12 months. Loss to follow-up was largely due to non-consent to further follow-up and non-response to mailing, though a small number of participants withdrew from the trial (reasons are provided in *Figure 14*).

Practice and patient characteristics

Practices were recruited from seven trusts in West Midlands North and practice deprivation scores spanned roughly 30–90th percentiles for England (i.e. towards the less deprived end of the spectrum). There was no difference in annual consultation prevalence for OA between the participating practices.

Under the minimisation algorithm, more practices were allocated to the control group and their total average practice list size was also higher than practices allocated to the intervention group (*Table 52*). Individual patients recruited from intervention and control practices had broadly similar characteristics (*Table 53*). In total, the average age of participants was 65 years and 57% were female. The largest difference was in the proportion of patients reporting the pain episode to be their first: 40% in the control arm versus 33% in the intervention arm. For the subgroup of 1035 study patients [644 (71%) in the control arm and 391 (78%) in the intervention arm] who consented to medical record review and had a baseline template pain score (recorded at GP consultation), the mean pain score was 6.33 (SD 2.04) in the intervention arm and 6.30 (SD 2.10) in the control arm.

TABLE 52 Cluster-level baseline characteristics according to study group

Characteristics of the participating general practices	Control ($N = 24$)	Intervention ($N = 20$)
PCT,[a] *n* (%)		
Region 1	6 (25.0)	6 (30.0)
Region 2	1 (4.2)	2 (10.0)
Region 3	5 (20.8)	4 (20.0)
Region 4	4 (16.7)	4 (20.0)
Region 5	1 (4.2)	1 (5.0)
Region 6	4 (16.7)	1 (5.0)
Region 7	3 (12.5)	2 (10.0)
Practice deprivation score,[a] median (IQR)	16.8 (11.1–37.9)	20.3 (11.2–29.3)
Total practice list size,[a,b] mean (SD)	7397 (4250)	5850 (2693)
Practice list size for ages ≥ 45 years, mean (SD)[c]	3519 (2170)	2736 (1244)
Consultation rate for OA in past 12 months (per 10,000 registered persons aged ≥ 45 years), mean (SD)	1590 (959)	1925 (746)

a Variable used in balance algorithm for randomisation.
b The sizes in the table relate to those used in the minimisation algorithm. Pre-audit figures are also available and the mean (SD) were: 7497 (4420) for the control group and 6008 (2664) for the intervention group.
c Correlation between total practice size and practice size for the ≥ 45 ages: 0.96.

Reproduced from Mallen *et al.*[300] © 2017 Mallen *et al.* This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 53 Participant baseline characteristics by study group

	Control (*N* = 911)	Intervention (*N* = 501)
Age (years), mean (SD)	65.4 (10.1)	65.6 (10.8)
Female, *n* (%)	509 (55.9)	290 (57.9)
White UK or European ethnicity, *n* (%)	888 (97.7)	492 (98.2)
Live alone, *n* (%)	182 (20.0)	108 (21.6)
Lack of emotional support, *n* (%)	41 (4.6)	36 (7.2)
Lack of support with daily tasks, *n* (%)	57 (6.3)	40 (8.0)
Currently in a paid job, *n* (%)	286 (31.4)	166 (33.1)
Time off work in previous 6 months, *n* (%)	102 (36.2)	65 (39.4)
Self-reported BMI (kg/m^2), mean (SD)	28.7 (5.6)	28.6 (5.1)
Previous/current smoker, *n* (%)	457 (50.8)	247 (49.3)
Daily/weekly alcohol consumption, *n* (%)	491 (54.4)	245 (49.0)
Area of pain consulted GP with, *n* (%)		
Neck	56 (6.2)	39 (7.8)
Shoulder	155 (17.0)	71 (14.2)
Elbow	20 (2.2)	20 (4.0)
Wrist or hand	121 (13.3)	77 (15.4)
Hip	224 (24.6)	130 (26.0)
Knee	496 (54.5)	282 (56.3)
Ankle or foot	126 (13.8)	63 (12.6)
First pain consultation episode, *n* (%)	355 (39.5)	163 (32.8)
Duration of complaint, *n* (%)		
< 3 months	219 (24.7)	118 (24.2)
3–12 months	261 (29.5)	110 (22.5)
1–5 years	269 (30.3)	152 (31.2)
> 5 years	137 (15.4)	108 (22.2)
Belief that pain started after accident/injury, *n* (%)	168 (18.6)	89 (18.1)
Comorbidity,[a] *n* (%)	666 (73.1)	370 (73.9)
Previous bone fracture, *n* (%)	375 (41.2)	215 (42.9)
Previous falls in past 12 months, *n* (%)	248 (27.4)	147 (29.5)
Template pain score,[b] mean (SD)	6.30 (2.10)	6.33 (2.04)

a Comorbidity: any of the following self-reported conditions – previous heart attack or stroke, angina, raised blood pressure, diabetes, circulation problems in legs, cancer, liver disease, kidney disease, asthma/bronchitis, deafness, eyesight problems.
b Values based on 1035 participants who consented to medical record review and who had baseline recorded template pain score (644 in the control group and 391 in the intervention group).

Reproduced from Mallen *et al.*[300] © 2017 Mallen *et al.* This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Clinical effectiveness

The results for analysis of the primary outcome measure (current pain intensity) including primary end point and secondary end point evaluations along with pre-specified ancillary analysis (sensitivity and subgroup analyses) are shown in *Table 54*. For the primary end point analysis there was a statistically significantly higher average pain score over the four follow-up time points in the intervention group than the control group (mean difference 0.33, 95% CI 0.05 to 0.61; effect size 0.16, 95% CI 0.02 to 0.29) (see *Table 54*). The difference was not uniform across individual time points; the largest difference of 0.50 was observed at 6 months follow-up. All three sensitivity analyses showed similar results and detailed analysis suggests that selection bias (GPs in the intervention arm tending to complete the template for more severe patients) was plausible, although not proven (data not shown). The estimates for the subgroup analyses showed a statistically non-significant trend for decreasing difference in pain scores between intervention

TABLE 54 Evaluation of the primary outcome measure (current pain intensity)[a]

Phase of statistical analysis	Post consultation[b] (control, *n* = 898; intervention, *n* = 493)	3 months (control, *n* = 701; intervention, *n* = 383)	6 months (control, *n* = 680; intervention, *n* = 374)	12 months (control, *n* = 644; intervention, *n* = 368)	Overall (control, *n* = 2923; intervention, *n* = 1618)
Main analysis (primary end point shaded)[c]					
Control, *n* [mean (SD)]	5.3 (2.7)	4.5 (2.8)	4.2 (2.9)	4.0 (3.0)	4.6 (2.9)
Intervention, *n* [mean (SD)]	5.7 (2.6)	5.0 (2.8)	4.8 (2.9)	4.3 (3.0)	5.0 (2.8)
Mean difference (95% CI)[d]	0.34 (0.01 to 0.66)	0.17 (–0.18 to 0.52)	0.50 (0.14 to 0.87)	0.32 (–0.10 to 0.74)	0.33[d] (0.05 to 0.61)
p-value	0.043	0.343	0.007	0.141	0.022
Effect size (95% CI)[e]	0.16 (0.00 to 0.32)	0.08 (–0.09 to 0.25)	0.24 (0.07 to 0.42)	0.15 (–0.05 to 0.36)	0.16 (0.02 to 0.29)
Sensitivity analysis (1)					
Mean difference (95% CI)	0.35 (0.00 to 0.69)	–0.02 (–0.39 to 0.35)	0.44 (0.06 to 0.82)	0.21 (–0.21 to 0.64)	0.25 (–0.02 to 0.52)
p-value	0.048	0.905	0.025	0.329	0.074
Control, intervention	595, 360	488, 292	451, 271	437, 254	1971, 1177
Sensitivity analysis (2)					
Mean difference (95% CI)	0.41 (0.04 to 0.78)	0.07 (–0.36 to 0.50)	0.36 (–0.12 to 0.85)	0.25 (–0.24 to 0.73)	0.37 (0.04 to 0.69)
p-value	0.029	0.753	0.137	0.324	0.030
Sensitivity analysis (3)					
Mean difference (95% CI)	0.39 (0.04 to 0.74)	0.19 (–0.19 to 0.56)	0.55 (0.16 to 0.95)	0.40 (–0.05 to 0.85)	0.37 (0.07 to 0.67)
p-value	0.030	0.335	0.006	0.083	0.015
Subgroup analysis (1)					
Mean difference (95% CI)	–0.15 (–0.44 to 0.13)	–0.35 (–0.67 to –0.03)	–0.14 (–0.47 to 0.19)	–0.32 (–0.70 to 0.06)	–0.21 (–0.46 to 0.03)
p-value	0.299	0.033	0.395	0.099	0.089

continued

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 54 Evaluation of the primary outcome measure (current pain intensity)[a] (*continued*)

Phase of statistical analysis	Post consultation[b] (control, *n* = 898; intervention, *n* = 493)	3 months (control, *n* = 701; intervention, *n* = 383)	6 months (control, *n* = 680; intervention, *n* = 374)	12 months (control, *n* = 644; intervention, *n* = 368)	Overall (control, *n* = 2923; intervention, *n* = 1618)
Subgroup analysis (2)					
Mean difference (95% CI)	0.03 (–0.12 to 0.19)	0.08 (–0.09 to 0.24)	0.04 (–0.13 to 0.21)	0.06 (–0.14 to 0.26)	0.05 (–0.08 to 0.18)
p-value	0.673	0.365	0.655	0.569	0.468

a Current pain intensity (0–10 numerical rating scale): 0 = no pain; 10 = pain as bad as could be. Values based on analysis of 4541 available data (*n* = 4541).
b Time between date of consultation and return of post-consultation questionnaire (days), median (IQR): control group 24 (17–35) days and intervention group 22 (16–33) days.
c Primary end point (overall pain intensity). The variance-partition for the random coefficients were as follows: ICC < 0.001 (for between-practice variation) and ICC = 0.591 (for between-individual variation).
d All analyses adjusted using GP practice and repeated measures as random-effects, and fixed-effect covariates at practice-level (as outlined in *Table 52*) and patient-level factors (age, sex and time between consultation and post-consultation response). Estimates at the individual follow-up time points were obtained by inclusion of interaction terms for study group by time point of assessment. Evaluation was by multilevel linear regression analysis with longitudinal random slope parameterisation.
e Estimated mean difference relative to SD of template pain score of 2.07.
Sensitivity analyses: (1) Estimates based on a subgroup of the study population who consented to medical record review and had a baseline recorded template pain score (as well as available post-consultation data on duration of pain and BMI): analysis based on 3148 available data by multilevel regression (as indicated in footnote d above with the addition of baseline pain template score, duration of pain and BMI as covariates); (2) 'per protocol' evaluation through CACE analysis of available data by instrumental variable analysis using two-step least squares regression; and (3) additional model adjustment for post-randomisation practice-level variables of number of patients mailed and number of patients who returned a completed post-consultation questionnaire (to account for selection bias in recruitment uptake to the study). Further, multiple imputation estimates for overall pain = 0.25 (95% CI –0.02, 0.53).
Subgroup analyses: multilevel linear regression analysis was carried out (as detailed in footnote d above) and included (1) group × age interaction term as the factor of interest, where age coefficients shown are based on units of 10 years [predicted mean difference in study population: 0.82 (45- to 54-year age group), 0.55 (55- to 64-year age group), 0.41 (65- to 74-year age group), 0.10 (≥ 75 years age-group)] and (2) group × pain (template pain score) interaction term as the factor of interest [predicted mean difference in study population: 0.17 (0–4 baseline pain template score), 0.37 (5–7 baseline pain template score), 0.61 (8–10 baseline pain template score)].
Adapted from Mallen *et al.*[300] © 2017 Mallen *et al.* This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

group and control group with increased age, but no strong evidence of an interaction with pain severity recorded by the GP at the point of care.

Secondary outcomes were consistent with the primary outcome measure in reflecting better outcomes as a whole for the control group than the intervention group (*Table 55*), including patient global assessment of change ($p = 0.015$).

The proportion of patients reporting that the GP asked irrelevant questions in the consultation was low and similar in both arms [41 (8.3%) in the intervention group and 50 (5.6%) in the control group] (*Table 56*). The proportion of patients not satisfied with the consultation was higher in the intervention group than in the control group [71 (14.5%) compared with 89 (9.9%), respectively].

General practitioner questionnaire

A total of 165 GP questionnaires were sent out and 82 were returned, giving a 50% response rate. The results are shown in *Table 57*. Most doctors in both arms reported an impact of the screening on consultation time at least sometimes but most did not find it difficult to incorporate in the consultation, with no statistically significant difference between interventions and controls. However, estimated

TABLE 55 Secondary outcome measures: mean (standard deviation), mean difference and *p*-value of difference, unless otherwise indicated

Outcome measures (range) (number analysed)	Post consultation	3 months	6 months	12 months	Overall time points
Worst pain (0–10 NRS) ($n = 4454$[a])					
Control	7.7 (1.9)	7.0 (2.4)	6.0 (2.8)	5.6 (3.1)	6.7 (2.7)
Intervention	7.8 (2.1)	7.2 (2.3)	6.6 (2.7)	5.8 (3.1)	7.0 (2.6)
Mean difference (95% CI)[b]	–0.01 (–0.30 to 0.28)	0.08 (–0.24 to 0.40)	0.54 (0.19 to 0.89)	0.52 (0.07 to 0.97)	0.15 (–0.09 to 0.40)
p-value	0.946	0.638	0.002	0.025	0.220
Average pain (0–10 NRS) ($n = 4433$)					
Control	6.4 (2.1)	5.7 (2.4)	5.0 (2.8)	4.7 (2.9)	5.5 (2.6)
Intervention	6.7 (2.2)	6.1 (2.4)	5.6 (2.7)	5.1 (2.9)	5.9 (2.6)
Mean difference (95% CI)[b]	0.25 (–0.03 to 0.53)	0.19 (–0.13 to 0.50)	0.54 (0.20 to 0.88)	0.48 (0.06 to 0.90)	0.30 (0.05 to 0.54)
p-value	0.081	0.240	0.002	0.026	0.017
Interference with daily activities (0–10 NRS) ($n = 4535$)					
Control	5.2 (2.8)	4.5 (3.0)	3.9 (3.1)	3.8 (3.1)	4.4 (3.0)
Intervention	5.4 (2.9)	5.4 (3.0)	4.6 (3.0)	4.3 (3.2)	4.8 (3.0)
Mean difference (95% CI)[b]	0.07 (–0.28 to 0.42)	0.11 (–0.27 to 0.48)	0.50 (0.12 to 0.89)	0.39 (–0.04 to 0.83)	0.20 (–0.10 to 0.51)
p-value	0.701	0.568	0.011	0.078	0.191
Interference with recreational activities (0–10 NRS) ($n = 4443$)					
Control	5.1 (3.1)	4.4 (3.2)	3.8 (3.2)	3.7 (3.2)	4.4 (3.2)
Intervention	5.5 (3.1)	4.7 (3.3)	4.5 (3.2)	4.0 (3.3)	4.8 (3.3)
Mean difference (95% CI)[b]	0.32 (–0.05 to 0.70)	0.12 (–0.28 to 0.52)	0.51 (0.09 to 0.92)	0.39 (–0.08 to 0.86)	0.32 (–0.01 to 0.65)
p-value	0.093	0.560	0.017	0.101	0.060
Interference with work (0–10 NRS) ($n = 4440$)					
Control	4.6 (3.1)	4.1 (3.1)	3.7 (3.1)	3.5 (3.2)	4.0 (3.2)
Intervention	5.1 (3.0)	4.4 (3.1)	4.3 (3.2)	3.9 (3.6)	4.5 (3.2)
Mean difference (95% CI)[b]	0.32 (–0.05 to 0.69)	0.09 (–0.30 to 0.48)	0.50 (0.10 to 0.91)	0.46 (0.00 to 0.92)	0.32 (–0.01 to 0.65)
p-value	0.092	0.655	0.015	0.049	0.059
Disability days in past month ($n = 4432$), *n* (%)					
Control					
0–3	348 (39.0)	326 (46.7)	361 (55.1)	355 (57.4)	1390 (48.6)
4–7	139 (15.6)	111 (15.9)	91 (13.9)	86 (13.9)	427 (14.9)
8–15	139 (15.6)	95 (13.6)	61 (9.3)	54 (8.7)	349 (12.2)
16+	266 (29.8)	166 (23.8)	142 (21.7)	123 (19.9)	697 (24.4)

continued

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 55 Secondary outcome measures: mean (standard deviation), mean difference and *p*-value of difference, unless otherwise indicated (*continued*)

Outcome measures (range) (number analysed)	Post consultation	3 months	6 months	12 months	Overall time points
Intervention					
0–3	158 (32.5)	156 (40.8)	166 (46.2)	170 (49.7)	650 (41.4)
4–7	88 (18.1)	69 (18.1)	65 (18.1)	54 (15.8)	276 (17.6)
8–15	73 (15.0)	57 (14.9)	50 (13.9)	44 (12.9)	224 (14.3)
16+	167 (34.4)	100 (26.2)	78 (21.7)	74 (21.6)	419 (26.7)
OR (95% CI)[b]	1.16 (0.69 to 1.96)	1.14 (0.64 to 2.04)	1.26 (0.66 to 2.42)	1.54 (0.68 to 3.48)	1.19 (0.76 to 1.87)
p-value	0.574	0.660	0.481	0.303	0.443
CPG pain subscale ($n = 4461$)					
Control	64.7 (19.4)	57.2 (22.9)	50.4 (26.6)	47.4 (28.1)	55.9 (25.0)
Intervention	67.1 (19.5)	61.0 (22.8)	56.7 (25.5)	50.3 (28.4)	59.6 (24.6)
Mean difference (95% CI)[b]	1.84 (–0.85 to 4.53)	1.41 (–1.54 to 4.36)	5.36 (2.16 to 8.56)	4.59 (0.51 to 8.66)	2.47 (0.05 to 4.88)
p-value	0.181	0.348	0.001	0.027	0.045
CPG disability subscale ($n = 4457$)					
Control	49.8 (28.1)	43.6 (29.5)	38.0 (30.1)	36.4 (30.5)	42.7 (29.9)
Intervention	53.2 (28.1)	46.7 (29.8)	44.7 (30.1)	40.0 (31.5)	46.8 (30.1)
Mean difference (95% CI)[b]	2.33 (–1.10 to 5.77)	1.10 (–2.55 to 4.76)	5.06 (1.26 to 8.86)	4.70 (0.32 to 9.08)	2.79 (–0.33 to 5.90)
p-value	0.183	0.554	0.009	0.036	0.079
CPG grade ($n = 4411$), *n* (%)					
Control					
I	141 (15.9)	212 (30.4)	266 (41.0)	278 (45.2)	897 (31.5)
II	280 (31.6)	175 (25.1)	133 (20.5)	118 (19.2)	706 (24.8)
III	163 (18.4)	115 (16.5)	99 (15.3)	85 (13.8)	462 (16.2)
IV	301 (34.0)	196 (28.1)	151 (23.3)	134 (21.8)	782 (27.5)
Intervention					
I	68 (14.1)	88 (23.1)	120 (33.4)	137 (40.2)	413 (26.4)
II	117 (24.2)	104 (27.3)	72 (20.1)	68 (19.9)	361 (23.1)
III	117 (24.2)	73 (19.2)	69 (19.2)	44 (12.9)	303 (19.4)
IV	181 (37.5)	116 (30.5)	98 (27.3)	92 (27.0)	487 (31.1)
OR (95% CI)[b]	2.18 (1.30 to 3.66)	1.24 (0.70 to 2.21)	1.51 (0.79 to 2.88)	1.37 (0.57 to 3.28)	1.72 (1.10 to 2.69)
p-value	0.003	0.461	0.215	0.478	0.016

TABLE 55 Secondary outcome measures: mean (standard deviation), mean difference and *p*-value of difference, unless otherwise indicated (*continued*)

Outcome measures (range) (number analysed)	Post consultation	3 months	6 months	12 months	Overall time points
Manikin – widespread pain (ACR) (*n* = 4401), *n* (%)					
Group 0	221 (24.3)	149 (21.6)	136 (21.4)	150 (24.9)	656 (23.1)
Group 1	128 (25.6)	91 (24.2)	94 (26.7)	82 (24.6)	395 (25.3)
OR (95% CI)[b]	0.95 (0.53 to 1.70)	0.97 (0.50 to 1.89)	1.28 (0.65 to 2.50)	1.01 (0.51 to 2.00)	1.03 (0.64 to 1.64)
p-value	0.865	0.938	0.471	0.972	0.917
Manikin – widespread pain (Manchester), *n* (%)					
Group 0	105 (11.6)	69 (10.0)	66 (10.4)	76 (12.6)	316 (11.1)
Group 1	69 (13.8)	62 (16.5)	63 (17.9)	44 (13.2)	238 (15.2)
OR (95% CI)[b]	1.09 (0.51 to 2.33)	2.77 (1.16 to 6.61)	3.94 (1.61 to 9.63)	1.24 (0.50 to 3.11)	1.77 (0.97 to 3.24)
p-value	0.816	0.022	0.003	0.641	0.063
WOMAC – physical functioning (0–32) (*n* = 4342)					
Control	12.5 (7.6)	11.3 (7.8)	10.6 (7.6)	10.6 (7.9)	11.4 (7.8)
Intervention	13.0 (7.7)	12.5 (7.6)	12.2 (7.9)	11.0 (8.1)	12.3 (7.9)
Mean difference (95% CI)[b]	0.42 (–0.41 to 1.25)	0.53 (–0.36 to 1.43)	0.70 (–0.23 to 1.64)	0.36 (–0.68 to 1.40)	0.49 (–0.27 to 1.26)
p-value	0.321	0.244	0.142	0.495	0.204
GAD-7 (0–21) (*n* = 4359)					
Control	5.1 (5.7)	4.6 (5.3)	4.6 (5.2)	4.6 (5.3)	4.8 (5.4)
Intervention	5.6 (5.8)	5.2 (5.6)	5.7 (5.9)	5.5 (6.1)	5.5 (5.8)
Mean difference (95% CI)[b]	0.18 (–0.49 to 0.84)	0.34 (–0.36 to 1.05)	0.44 (–0.28 to 1.16)	0.61 (–0.14 to 1.37)	0.34 (–0.27 to 0.94)
p-value	0.603	0.341	0.230	0.112	0.272
PHQ-8 (0–24) (*n* = 4376)					
Control	6.0 (6.0)	5.2 (5.7)	5.3 (5.7)	5.4 (6.0)	5.5 (5.8)
Intervention	6.4 (6.1)	6.0 (6.1)	6.6 (6.3)	6.0 (6.1)	6.3 (6.1)
Mean difference (95% CI)[b]	0.30 (–0.40 to 1.00)	0.52 (–0.22 to 1.26)	0.74 (–0.02 to 1.49)	0.36 (–0.41 to 1.14)	0.44 (–0.21 to 1.09)
p-value	0.402	0.172	0.055	0.360	0.181
SF-12 PCS (0–100) (*n* = 4263)					
Control	36.0 (11.1)	37.9 (11.4)	39.3 (11.8)	39.1 (11.9)	37.9 (11.6)
Intervention	35.5 (10.5)	36.3 (10.8)	36.3 (11.3)	38.1 (11.6)	36.4 (11.0)
Mean difference (95% CI)[b]	0.24 (–1.07 to 1.55)	–0.23 (–1.63 to 1.17)	–1.77 (–3.22 to –0.32)	–0.66 (–2.25 to 0.93)	–0.33 (–1.54 to 0.89)
p-value	0.717	0.749	0.017	0.419	0.598

continued

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 55 Secondary outcome measures: mean (standard deviation), mean difference and *p*-value of difference, unless otherwise indicated (*continued*)

Outcome measures (range) (number analysed)	Post consultation	3 months	6 months	12 months	Overall time points
SF-12 PCS (0–100) ($n = 4263$)					
Control	49.9 (11.4)	49.6 (11.5)	49.0 (11.7)	49.2 (11.3)	49.5 (11.5)
Intervention	49.1 (11.2)	48.4 (11.5)	47.6 (12.0)	48.8 (11.6)	48.5 (11.6)
Mean difference (95% CI)[b]	–0.61 (–1.98 to 0.76)	–0.79 (–2.26 to 0.69)	–0.12 (–1.62 to 1.39)	–0.32 (–1.88 to 1.25)	–0.50 (–1.73 to 0.72)
p-value	0.383	0.295	0.878	0.691	0.418
Perceived change ($n = 2977$), *n* (%)					
Control					
Completely recovered	–	26 (3.8)	43 (6.8)	63 (10.5)	132 (6.9)
Much improved	–	98 (14.3)	135 (21.4)	141 (23.4)	374 (19.5)
Improved	–	177 (25.8)	140 (22.2)	104 (17.3)	421 (22.0)
No change	–	237 (34.6)	166 (26.4)	154 (25.6)	557 (29.0)
Worse	–	121 (17.6)	112 (17.8)	115 (19.1)	348 (18.1)
Much worse	–	27 (3.9)	34 (5.4)	25 (4.2)	86 (4.5)
Intervention	–				
Completely recovered	–	13 (3.5)	21 (6.0)	24 (7.2)	58 (5.5)
Much better	–	52 (13.9)	47 (13.4)	69 (20.7)	168 (15.9)
Somewhat better	–	70 (18.7)	62 (17.6)	57 (17.1)	189 (17.9)
No change	–	139 (37.2)	109 (31.0)	101 (30.3)	349 (33.0)
Worse	–	80 (21.4)	89 (25.3)	63 (18.9)	232 (21.9)
Much worse	–	20 (5.4)	24 (6.8)	19 (5.7)	63 (6.0)
OR (95% CI)[b]	–	0.66 (0.34 to 1.30)	1.81 (0.95 to 3.45)	1.88 (1.02 to 3.46)	1.37 (0.84 to 2.24)
p-value	–	0.232	0.073	0.044	0.200

ACR, American College of Rheumatology criteria; CPG, Chronic Pain Grade; GAD, Generalised Anxiety Disorder.
a Total number of available longitudinal data that were analysed (in each case).
b All analyses adjusted using GP practice and repeated-measures as cluster-level random-effects and fixed-effect covariates at practice level (as outlined in *Table 54*) and patient level (age, sex and time between consultation and post consultation response). Analysis was carried out by linear-mixed model with unstructured covariance matrix.

Adapted from Mallen *et al.*[300] © 2017 Mallen *et al.* This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

TABLE 56 Acceptability and fidelity of screening from participants' post-consultation questionnaires

Items on patient questionnaire	Control ($N = 911$), *n* (%)	Intervention ($N = 501$), *n* (%)	Statistical significance
Doctor asked about			
How long you have had your pain	696 (76.4)	368 (73.5)	$p = 0.219$[a]
The intensity of your pain	674 (74.0)	348 (69.5)	$p = 0.069$
How your pain interferes with daily activities	410 (45.0)	236 (47.1)	$p = 0.448$
Your mood	90 (9.9)	157 (31.3)	$p < 0.001$
Doctor asked irrelevant questions			
No	743 (82.6)	391 (79.1)	
Unsure	106 (11.8)	62 (12.6)	
Yes	50 (5.6)	41 (8.3)	$p = 0.117$
Irrelevant questions on mood, *n*[b]	3	7	
Satisfied with consultation			
Yes	676 (75.5)	342 (69.8)	
Unsure	130 (14.5)	77 (15.7)	
No	89 (9.9)	71 (14.5)	$p = 0.025$

a χ^2 test.
b From free-text responses to an open-ended question on irrelevant questions, completed by 29 and 27 patients in the control and intervention groups, respectively.

Reproduced from Mallen *et al.*[300] © 2017 Mallen *et al.* This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

TABLE 57 Participating GPs' perceived impact of screening questions on the consultation

Items on patient questionnaire	Control ($N = 51$), *n* (%)	Intervention ($N = 31$), *n* (%)	Statistical significance
Ease of incorporation into the consultation			
Very easy	11 (21.6)	5 (16.1)	
Quite easy	19 (37.3)	9 (29.0)	
Neither easy nor difficult	14 (27.5)	10 (32.3)	
Quite difficult	5 (9.8)	7 (22.6)	
Very difficult	2 (3.9)	0 (0.0)	$p = 0.194$[1]
Impact on consultation time			
Always	5 (9.8)	2 (6.5)	
Often	3 (5.9)	8 (25.8)	
Sometimes	18 (35.3)	14 (45.2)	
Rarely	18 (35.3)	6 (19.4)	
Never	7 (13.7)	1 (3.2)	$p = 0.014$[2]

continued

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 57 Participating GPs' perceived impact of screening questions on the consultation (*continued*)

Items on patient questionnaire	Control ($N = 51$), *n* (%)	Intervention ($N = 31$), *n* (%)	Statistical significance
On average, how much additional time			
No additional time	2 (4.5)	1 (3.3)	
Less than 1 minute	19 (43.2)	5 (16.7)	
1–2 minutes	17 (38.6)	14 (46.7)	
2–3 minutes	6 (13.6)	7 (23.3)	
3–4 minutes	0 (13.6)	1 (3.3)	
4–5 minutes	0 (0.0)	2 (6.7)	
More than 5 minutes	0 (0.0)	0 (0.0)	$p = 0.008$[3]
Impact on doctor–patient communication			
Always	1 (2.0)	1 (3.3)	
Often	4 (8.0)	1 (3.3)	
Sometimes	11 (22.0)	16 (53.3)	
Rarely	20 (40.0)	7 (23.3)	
Never	14 (28.0)	5 (16.7)	$p = 0.033$[4]
Nature of impact on doctor–patient communication			
Positive	8 (26.7)	13 (56.5)	
Negative	1 (3.3)	1 (4.3)	
Both positive and negative	14 (46.7)	3 (13.0)	
Do not know	7 (23.3)	6 (26.1)	$p = 0.028$[5]
Impact on patient management			
Always	1 (2.0)	1 (3.2)	
Often	0 (0.0)	2 (6.5)	
Sometimes	7 (13.7)	13 (41.9)	
Rarely	22 (43.1)	7 (22.6)	
Never	21 (41.2)	8 (25.8)	$p = 0.005$[6]

additional time for screening was higher among intervention than control doctors. A statistically significantly higher proportion of doctors from the intervention practices reported an impact on patient management and communication with the patient, and reported feeling that this was positive.

To calculate chi-squared test for trend statistics, categories were merged and the following six categories, as numbered, were used in the analysis: (1) 'very/quite easy', 'neither easy nor difficult', 'very/quite difficult'; (2) 'always/often', 'sometimes', 'rarely/never'; (3) '< 1 minute', '1–2 minutes', '> 2 minutes' (test excluded responders who reported 'never' to the question on 'impact on consultation time'); (4) 'sometimes/often/always', 'rarely', 'never'; (5) 'positive', 'not positive' (test excluded responders who reported 'never' to the question on 'impact on doctor–patient communication'); (6) 'always/often/sometimes', 'rarely', 'never'.

General practitioner interview study recruitment

Responses to the request (contained in the GP survey questionnaire) to take part in a further interview phase of the study are summarised in *Table 58*. Out of the 82 GPs who returned questionnaires, six declined any follow-up contact to discuss their experiences of the study. This gave a total of 76 GPs from 26 practices: 15 practices from the intervention arm and 11 from the control. On contact, a further 13 GPs declined to be interviewed and three were lost to follow-up either through retirement or changing jobs. It was not possible to make contact with a further 25 GPs despite frequent attempts based on the details provided in the questionnaire, leaving 35 GPs available for interview.

A number of contact strategies were adopted. When GPs had given their e-mail addresses, these were used in the first instance. However, only one response was received from the initial mailing to 15 individuals. As a consequence, the decision was taken to make follow-up contact through the practice managers, and to adopt that approach in the first instance for other potential participants. The first two attempts at contact with the practice managers were via e-mail and when there was no response this was followed by a telephone call. The average number of contacts needed to obtain an initial response from the practices was three. This response in all instances involved the practice manager agreeing to take the matter to the GPs and a series of communications then often followed to remind them again of the study and to arrange the actual interviews. Twenty-five GPs were interviewed in total. Twelve were interviewed face-to-face; these included three who had not completed the questionnaire but who joined with four of their practice partners when they were interviewed as a group, two other GPs who were interviewed together, and three GPs seen on their own. The other 13 GPs had interviews conducted by telephone. Interviews lasted between 15 and 40 minutes.

Despite these challenges, data saturation was reached by the tenth interview in both arms of the study. Data collection continued for a further four participants in the control arm and two in the intervention arm with whom interviews had already been arranged. This confirmed that no new data were emerging and data collection then ceased. Ongoing liaison with the remaining practices to arrange interviews was cancelled.

This process gave a total of 25 interviews. Data saturation took a period of 7 months and ended in March 2013.

General practitioner interview study: emergent themes

Results are presented under the headings of the three major themes identified through the analysis: participation in the trial and using the screening tool, OA and accompanying psychological problems, and the doctor–patient relationship.

TABLE 58 Summary of responses to request for interview (GPs)

Stage in recruitment of GPs for interview	Number of GPs
Questionnaires sent	165
Questionnaires returned	82
Consent to follow-up	76
Declined at follow-up	13
Lost to follow-up	3
No response to follow-up	25
Available for interview	35
Agreed to interview without questionnaire	3
Total interviewed	25 (12 from control practices; 13 from intervention)

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

There was substantial overlap in these themes.

Participating in the trial and using the screening tool

General practitioners were asked whether or not the training on participation provided by the research team could be improved in any way, and the uniform response was that it had been very helpful and appropriate with no suggestions of any areas that needed to be strengthened for future research:

> *I can't really fault the way it works, I think I'm just grateful for the fact it was done quite simply. It flagged up when the relevant coding or diagnosis was put in, and helped me obviously think, and just a good reminder.*
>
> *34:85 C*

Although the account of recruitment given in the previous section illustrates difficulties in recruiting GPs to the interview study, analysis of both the questionnaires and the interviews indicate that most GPs who took part in the qualitative studies did not find taking part in the trial itself problematic. Nevertheless, although GPs in the control arm reported that the study required < 1 minute of additional time in the consultation, those in the intervention arm reported that it took between 1 and 3 minutes more.

There were some reports of excluding patients who were eligible for the trial in terms of the protocol and a number of reasons were given. The most common of these was that the template fired after the patient had left the room owing to a consultation style in which the GP entered computer data after the consultation. One GP reported that:

> *On a few occasions I had to chase patients down the corridor to bring them back and ask the questions.*
>
> *18:5 I questionnaire free-text response*

There were also reports of not including patients because of time pressures, as well as some reports of patients declining. One GP wrote on the questionnaire that the patient declined because of 'anxiety', but it was not possible to explore this in more detail as the GP gave no further information and declined a follow-up interview. Another GP noted that a patient declined because they felt they could not complete the questionnaires and, again, no further information was given and the GP declined to be interviewed.

When these issues were followed up with GPs who were interviewed, age emerged as a reason:

> *I believe one or two patients may not have been screened because of very complex ageing situations.*
>
> *17:5 I*

In general, GPs interviewed found screening tools to be of some use, particularly in referring patients to other resources, but felt they were no substitute for their own clinical judgement:

> *Clearly we have to do the PHQ for QOF, for patients with anxiety and depression that needs to be done. Our local EWB [Elderly Wellbeing] team . . . will not see a patient without it being filled in so it kind of . . . it gets to be something that you just do . . . [it] helps to expand on questions, you say 'well I can see on this last question you've marked it as two, can you tell me a little bit about that?' So it can sometimes open up the conversation. It's more useful as a tool so that I can record, and for further services to be involved. I don't need it for my skills in diagnosing depression, but it kind of adds to it really.*
>
> *5:20 I*

Although some GPs interviewed did not find the screening template intrusive, others did. One GP reported not using it because of:

> *. . . knowing the patient had enough on their plate with other issues.*
>
> *4:12 C questionnaire free-text response*

Others found it formulaic and awkward, even a nuisance:

> *I think we feel they are of a nuisance value. If this were made routine I don't think it would be difficult to incorporate, but I think we're generally trying to sort of speed up rather than give ourselves more to do.*
>
> *1:23 I*

Moreover, it was screening on a topic for which they would avoid using a case finding approach anyway:

> *I don't probe unless they say something about it because sometimes I feel it's very difficult to put that in their mind that they may be depressed or, if I feel that they are depressed I ask them, if I don't feel they are, then I don't.*
>
> *6:21 C*

However, for others it was helpful:

> *In general coding gave opening to discuss 'emotional psychological' impact of pain on patient's situation.*
>
> *17:5 I*

For the majority, high workloads and competing demands were key factors in not routinely case finding for anxiety and depression in this patient group:

> *The whole problem is there's so many things you've got to do, you've got to prioritise and we've got all the targets we've got to do, and all the QOF stuff . . . that there's very little time to go searching for things, so it tends to have to be opportunistic.*
>
> *34:81 C*

For some, however, taking part in the study raised awareness:

> *I think we always try to be patient centred anyway, but I think perhaps in a way, patients and doctors fall into a trap of discussing things in the same way, time after time, and actually having you know, entering into the study perhaps makes you focus on it slightly differently.*
>
> *5:17 I*

This increased awareness extended to some in the control group, even though it was only the numeric pain rating item that was additional to the consultation in that arm:

> *It's helped me to concentrate more on patients' needs than I would do, like with the elderly with arthritis, like with the wear and tear and things.*
>
> *6:21 C*

Although these views represent saturation in terms of the theme, the poor response to the interview study must be taken into account when judging the representativeness of these views of GPs as a whole. Out of the 165 questionnaires issued, just 35 GPs were available for interview (i.e. 21% of those mailed). The extent to which OA and its concomitant mental health challenges are seen as pressing within primary care must be considered with caution.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Osteoarthritis and accompanying psychological problems

The presence of depression and/or anxiety was seen as common:

> *I wouldn't say very common but certainly fairly common, I would say maybe 25% or so have got something psychological going on at the same time.*
>
> *5:18 I*

Indeed, some GPs considered depression and anxiety a normal part of having OA, and felt it was viewed in this way by patients:

> *I don't think they see it as a separate problem. I think if they're down because of their pain, I don't think most of them want to be given a label of psychological problems. I think they would understand it is almost to be expected.*
>
> *1: 23 I*

Despite such prevalence, there was awareness that mental health issues were not always appropriately recognised or addressed:

> *It's a major part of it. I think it's something that is potentially slightly under diagnosed or slightly overlooked.*
>
> *34:85 C*

Prioritisation of the physical over the mental and of depression over anxiety

All interviewees recognised the potential impact of OA on overall well-being and the complexity involved. However, GPs talked about prioritising the physical over the mental:

> *It's still difficult to untie it from the physical symptoms, isn't it? Because if they say 'I've got poor sleep', well 'that's because you've got pain'; they've got poor appetite, that's because I've got you on non-steroidals.*
>
> *F4: focus group C*

Among younger patients, maintaining capacity for, or getting back to, work emerged as a key priority, whereas for older patients it was seen as a more complex issue:

> *With most elderly patients presenting with joint issues, one of the things we'd always be keen to find out is how their condition was impacting at home, say for example mobility issues, bathing issues, falls that type of thing. The psychological aspects, I'd say that those get triggered with the more intractable patients with more severe disease, or perhaps those in which pain analgesia hasn't been easy to achieve. Always does sort of make you think about depression and anxiety as issues.*
>
> *14:28 C*

Participants also reported that their patients prioritised the physical, as this interview extract in *Box 3* highlights.

This was echoed in another interview when the GP participant was asked to talk through a typical response to a patient with depression and joint pain:

> *Well, if somebody presented we would always fill in a PHQ-9 . . . and then you would talk through them what the options are . . . 'do you want some medication for this?' A lot of people will immediately say 'no, I don't want any more tablets' and you say 'OK, there are counselling options' and mostly they don't want that either. The impact clinic or the pain clinic, they will take that option.*
>
> *34:81 C*

BOX 3 Interview extract (34:82 C)

I: *So, in your experience then how common are psychological or psychosocial problems in patients with joint pain or OA would you say?*

IV: *Well, if you go fishing they're very common. They don't always present to start with.*

I: *So when you say they don't always present, is that because patients tend not to self-report, or something else?*

IV: *Yes, they tend not to self-report, and they feel more comfortable talking about the physical symptoms.*

I, interviewer; IV, participant.

Again, this response was considered to be particularly common in older people:

> *Often I find elderly patients tend to rarely mention anything to do with low mood. They'll admit there is a problem to do with that when things are really, really desperate, or really bad. But some patients were raised during an era when even just making a trip to the doctors you had to be really debilitated before you got some help.*
>
> *34:85 C*

When it came to case-finding, depression was generally prioritised over anxiety, which tended to be seen as a one-off issue at most, rather than anything that might accumulate over time and, again, both came after the physical:

> *I wouldn't identify anxiety as being a factor in patients with osteoarthritis. If it does rear its head, it can usually be addressed fairly early on in the disease process by reassurance about what's going on. Whereas if we're not able to get adequate pain control and people's mood deteriorates, that makes it more difficult to control the pain.*
>
> *23–71 I*

As previously highlighted, this notion of prioritisation was contextualised within the constraints of the consultation and managing complex health needs in that time frame, but it was also very much framed by how the Quality Outcomes Framework was perceived to determine the valid use of time and resources in the face of such constraints:

> *We've got so many other things to do, is the problem . . . clearly if arthritis became a QOF target, anxiety, depression, the same sort of questions that you've got for diabetes and ischaemic heart disease, then that would be valid, wouldn't it?*
>
> *34:81 C*

The doctor–patient relationship

For most participants, the complex relationship between OA, depression and/or anxiety was reflected in the quality of the doctor–patient relationship, which emerged as key in being able to appropriately and sensitively address complex physical and mental health issues, especially when it came to older patients (*Box 4*).

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

BOX 4 Interview extract (34:85 C)

IV: *I think for the elderly patient, it's much easier when they know that you're an approachable, good communicator. They can understand you; you approach the consultation in a way that allows them also to say something to you, so time investment in that often helps. And I know it's a very difficult thing sometimes, to invest time because of pressures of work etc.*

I: *Yes. And of course that level of communication clearly affects the quality of treatment that patients receive?*

IV: *Yes, absolutely. And they're compliant as well. I think if they're trusting you and they're happy with the care they get they're likely to be more compliant and if they're not happy with treatment they're likely to communicate it to you instead of keeping quiet about it.*

I: *And do you find those sorts of issues are any different for patients with anxiety as opposed to patients with depression, or is it more or less the same?*

IV: *I think it's pretty more or less the same. I think sometimes in depression and anxiety there's the tendency to feel a bit withdrawn. Maybe they don't want to be opening up a lot, because they're either scared or you feel like you can't be bothered to say anything, or you're just so down in a mood that you just have lost the will, or lost the zest, for life. So I think it applies to both groups of patients.*

I, interviewer; IV, participant.

Conclusions to the qualitative study of general practitioners

Difficulties in recruiting to the interview study reflect current resource pressures and demands on GPs. Among those who responded, participation in the trial appears to have had little impact on the consultation in terms of time and to have raised their awareness of mental health issues among patients with OA. However, although case-finding was seen to be of some use, the use of screening tools was considered limited and, at best, an adjunct to clinical judgement. There was a sense that low mood in patients with OA was 'normal', particularly in older patients, and that GPs prioritised the physical over the mental. These views were reinforced by the absence of OA within the QOF, with GPs indicating that different QOF priorities might change their approach. When mental health was addressed, GPs prioritised depression over anxiety, with the depression associated with OA generally seen as normal but the importance of detecting and managing anxiety generally unrecognised. GPs also reported a paucity of non-pharmacological resources available for those patients in whom mental health difficulties were identified. Continuity of care was considered very important in the management of patients with long-term conditions such as OA, and a high-quality doctor–patient relationship was seen as essential to the effective management of consequential and coexisting mental health issues, particularly among older patients with multimorbidities. The use of guidelines in the management of this patient population was seen as helpful and the need for continuing professional training and updating was recognised.

Discussion

The primary objective was to evaluate the clinical effectiveness of introducing a point-of-care electronic template to prompt GPs to ask brief questions about anxiety and depression symptoms in patients consulting for clinical OA. Specifically, we hypothesised that patients in the intervention practices that screened for anxiety and depression symptoms in the GP consultation would show greater improvements in current pain intensity and pain interference with daily activity over the 12 months following their consultation than patients in the control practices.

Principal findings

This pragmatic cluster RCT in UK primary care provides evidence against a beneficial effect of introducing a template to screen patients consulting with OA for the presence of anxiety and depression on patient-reported OA outcomes.

The further observation of worse pain and functional outcomes across the 12 months following the index consultation among trial participants in the intervention practices could be explained, at least in part, by selection bias, although patient dissatisfaction with the consultation may also make a modest negative contribution.

Summary comparison with previous findings

This is the first such trial in OA to our knowledge and, in particular, was novel in tackling the issue of comorbid anxiety as well as depression in OA patients. The trial was highly pragmatic in nature in that it simply introduced the screening as a potential prompt for action by GPs without specifying what that action should be beyond highlighting the existence of NICE guidelines. Therefore, the trial took place in the context of existing mental health services and allowed clinician judgement on how to respond to screening questionnaires.

The results were generally consistent with limited effects of point-of-care reminders[132] and also with the idea that screening without addressing other important barriers (e.g. acceptability, referring on, having good mental health services in place) is unlikely to have any effect.[313]

There was no strong evidence that the intervention posed major problems for the patients or practitioners, although the themes that emerged from the nested qualitative study do highlight the variable views of GPs on the value of such screening and appear consistent with Maxwell *et al.*[314] Some of these concerns may have been averted by selective application of the template by GPs, hence giving rise to the potential source of selection bias in the study.

The hypothesis under test

Patients with painful, disabling OA constitute a 'high-risk' group for distress, anxiety and depression disorders.[35,37,315] Although many of the factors associated with the future course of OA are not modifiable at the point of care (e.g. age, sex, symptom duration, severity of underlying structural changes to the joint[316–319]) comorbid depression and anxiety are an exception as they are potentially treatable. They are related to future course,[320,321] treatment response[36] and health-care use,[322] and show a reciprocal relationship with pain and functional outcomes.[36,323,324] The initial idea for POST drew on evidence from clinical trials of collaborative care approaches that had used psychological therapies and medication for major depressive disorder in patients with OA to modify general and OA-specific clinical outcomes, and shown beneficial effects on pain intensity, pain-related function and QoL sustained to 12 months.[18,279,325] A novel component of POST was that it also considered anxiety, based on the idea that physical function benefits might accrue from effective management of comorbid anxiety disorders in patients with persistent painful disorders.[326,327]

The POST focused on screening for depression and anxiety in patients consulting about joint pain and OA. Current NICE guidelines for OA, updated since the inception of the trial, remain ambiguous on the matter of screening, recommending that patients be assessed for the effect of OA on mood, specifically including 'screen for depression' as a topic 'worth assessing', but acknowledging that it may not be of concern for every patient.[153] Evidence and guidelines that informed the design of POST encouraged practitioners to be alert to depression in patients with chronic physical conditions[287] and to consider asking two short screening questions of patients who they suspect of having depression.[328] Physician behaviour can be expected to influence the expression of emotional cues and concerns by patients and there is evidence that both elicitation and recognition is highly variable between practitioners.[329] Closed questions regarding psychosocial issues may facilitate the expression of emotional cues and concerns.[330,331] In this alternative perspective of screening, in which the purpose is to facilitate a more holistic assessment in order to

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

improve pain and functional outcomes, POST has investigated whether or not there are beneficial effects of raising the issue of anxiety and depression within the OA consultation and we hypothesised that these need not be restricted to the relatively small minority of patients ultimately diagnosed with anxiety or depressive disorder who access and receive high-quality mental health care (the 'screen-diagnose-treat' pathway). The recognition of subthreshold anxiety and depression symptoms, which are more common than clinically diagnosed anxiety and depression, and still associated with less favourable pain and function outcomes, could 'open the door to a dialogue with clinicians who can then determine which unmet needs have contributed to distress'.[313] This could include exploring causes (e.g. poorly controlled pain,[332] sleep disturbance,[333] inadequate social support[38]) as well as prompting greater use of pain management and functional rehabilitation options, such as referral to physiotherapy for supervised exercise that is effective for pain,[92] function[92] and mental health,[334,335] but are typically under-utilised.[101,336]

This thinking informed and justified our approach, which was to test the effect of introducing screening questions per se into the general practice consultation with OA patients without specifying or providing additional treatments or resources beyond those which the GP would then choose to use in usual practice. The results of POST provide no evidence of improved patient-reported pain and function to support this strategy and intervention as a means to manage OA in general practice.

Contributing evidence to the debate about screening for psychological distress

POST found a statistically significant negative effect on pain in the intervention practices. Selection bias is one plausible explanation for this, but the result may reflect a negative effect of screening in these circumstances. This links to an evolving debate in the literature about the benefits and harms of routine screening for depression in high-risk groups. In both unselected primary care populations and special populations at high-risk of depression, several recent systematic reviews by Meijet *et al.*,[337] Thombs *et al.*[338,339] and others[340] have highlighted the lack of direct evidence from appropriately designed clinical trials on the effects of implementing routine depression screening versus no screening, either alone or in the context of accessible, good-quality mental health care. Measured against this standard of evidence, the recommendation of routine depression screening in general,[339] and in high-risk patients with diabetes,[341] cancer[342] and coronary heart disease,[343] has been criticised as premature. Furthermore, these and other commentators[344] have expressed concern over the inefficiency of screening, the diversion of scarce resources from other clinical priorities, the potential for wasteful and unnecessary exposure to common side effects of antidepressant medications, particularly in patients with mild, self-limiting depressive symptoms, and possible nocebo effects and stigma from overt labelling. Counter arguments have stressed that, nevertheless, much has been learned from previous studies,[313] that the efficiency of screening improves if appropriately targeting high-risk groups[345] and both efficiency and acceptability may be helped by shifting the focus from 'depression' to a broader concept of distress.[313] Although modest increases in the rate of new depression diagnoses and in antidepressant prescriptions were observed following introduction of financial incentives for depression screening in patients with diabetes and coronary heart disease in UK primary care in 2006,[346,347] the appropriateness of these changes and their relation to patient outcomes remains unclear. POST adds importantly to this debate by showing no benefit from screening in older patients consulting about joint pain or OA. Although a negative effect of screening cannot be ruled out, bias is a more likely explanation. However, failure to find a true beneficial effect because of selection bias is improbable.

Strengths

This trial is one of the few that actually addresses an evidence gap that is still being highlighted in the current literature.[348] The screening was done by the GP in the consultation at the point of care and not by means of a remote questionnaire or a screening instrument applied separately to the consultation.[284] There was secure cluster allocation and minimal cluster attrition. Sample size was achieved and by the required deadline.

Limitations

There was cluster imbalance that resulted from variable list sizes of the randomised practices. However, this was unlikely to have had an adverse effect on statistical power and was unlikely to have introduced bias per se.

Eligibility criteria were broad. We have discussed, in *Chapter 3*, the potential problems in relation to the MOSAICS trial raised by the variety of patients who consult with one of the Read codes for joint pain or OA used to define initial eligibility in both MOSAICS and POST, a substantial proportion of whom may have minor symptoms or non-OA diagnoses that may dilute the effect of screening because they are not 'high risk' enough to benefit or to show improvements in pain and disability. In addition, in POST, the failure to exclude all those already known by the GP to be anxious or depressed may be regarded as a limitation of a screening programme, although the purpose of the intervention (to shape the GP consultation with the knowledge that a patient is anxious or depressed) could equally apply to a 'reminder' as to a 'first ever' structured prompt.

The qualitative study raises issues of fidelity of the intervention, namely the extent to which questions were not asked in the consultation but inferred afterwards owing to entering data at end of the consultation. This problem may have been selectively greater in the intervention than the control practices because of the nature and length of the screening tool.

This was basic binary screening, using a small number of questions to establish if a state of depression or anxiety was or was not present. The depression questions had been validated before for such use but there was no previous validation of this approach to anxiety screening.

Selection bias is plausible.[187] Cluster randomisation was chosen because the viability of individual randomisation was not realistic in the context of point-of-care screening. Prior identification of eligible participating patients could have been done, as per the method adopted for the MOSAICS trial; however, the effect would have been to markedly increase numbers in trial but with the majority not receiving the intervention and, therefore, even less likely to have an effect.

Implications

Whether or not the negative result was real or due to bias, it is clear that, at the very least, screening for depression and anxiety at point of care did not improve pain or function in patients aged $\geq$ 45 years who were consulting with joint pain or OA. The implication is that it is not possible to recommend this intervention for general practice.

This is important evidence to contribute to a debate that has been evolving since the trial was planned and designed, in which the value, role and usefulness of short systematic screening for symptoms of psychological distress in primary care are being questioned, despite acceptance that such symptoms are strongly associated with the health status of persons with long-term conditions such as OA.

Conclusion

Encouraging GPs to routinely ask screening questions for anxiety and depression of patients consulting for painful OA (and then follow guideline recommended care for OA and mental health) has no appreciable benefit on patient-reported pain and functional outcomes.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Chapter 6 Conclusions

The conclusions from the individual workstreams of the programme have been presented in the relevant chapters and the overall conclusions from the programme have been presented in *Chapter 1*.

In this chapter we summarise briefly again the picture that emerges from the programme as a whole and possible implications for research and practice (see *The results of the programme* and *Recommendations for future work*), before presenting comparative data on the different populations studied in the programme (see *Overview of the workstream populations*) and describing some other outputs that emerged from the programme as a whole (see *Outputs from delivery and dissemination of the programme*).

The results of the programme

Summary of the main findings

This programme has investigated primary care for OA patients.

Across four workstreams, the programme has investigated:

- the potential for primary care to have an impact on the prevention of progression of OA in the community
- how and whether or not the delivery of NICE core guidance for OA care in general practice can be improved, and whether or not delivery of NICE core guidance in general practice improves pain and disability among OA patients in primary care
- whether or not adherence to exercise and physical activity can be improved by changes to physiotherapy practice among patients with knee OA
- whether or not screening older patients who consult about joint pain and OA for depression and anxiety leads to improvements in their pain and disability.

The conclusions are that (1) on the basis of available evidence about efficacy in tackling prognostic factors for poor outcome in patients with OA, there is substantial potential for primary care to influence the frequency and severity of OA in the community and (2) it is feasible, acceptable and achievable to (i) improve quality of care in general practice for OA, in particular the delivery of NICE OA core guidance, through GP and practice nurse training and through use of quality-of-care templates inserted in GP records, (ii) deliver enhanced physiotherapy to support long-term adherence to exercise and physical activity in patients with knee OA and (iii) identify the substantial numbers of patients presenting with joint pain in primary care who also have symptoms of depression or anxiety.

However, the improved delivery of NICE core guidance by GPs and nurses achieved in this programme did not result in statistically significant improvements in pain and disability in patients with OA compared with usual general practice care, although there was evidence of more efficient prescribing and patient satisfaction with their care. In addition, enhanced physiotherapy approaches to sustaining exercise and physical activity in patients with knee OA did not result in statistically significant improvements in pain and disability compared with UC, and screening for depression and anxiety in patients consulting about joint pain did not improve their pain and disability.

Main implications for practice

- Improved implementation of NICE OA core guidance can be achieved in general practice through patient-generated information, training programmes for health-care professionals, and structured quality-of-care reminders.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

- New ways need to be found to sustain the known effectiveness of physiotherapy-led exercise and physical activity interventions for persons with knee OA which declines shortly after contact with the physiotherapist has ended.
- Screening patients who consult with joint pain in general practice for symptoms of depression or anxiety cannot be recommended as a policy for improving pain and disability in these patients.

1. Possible explanations for lack of effect of successful implementation of NICE guidance in general practice on patient clinical outcomes, and their implications for practice.

    i. The content and level of the intervention achieved through GP consultation and practice nurse-led clinics may be different from that carried out in the efficacy studies on which NICE core guidance for OA was based.

        – The implication is that, as improvements to the quality of general practice by delivering on NICE core guidance were insufficient to have an impact on patients' pain and disability, more intensive or specialised interventions, such as physiotherapy-led exercise, may be routinely required in primary care.
        – Different models of primary care for patients with OA, which can deliver more intensive evidence-based treatment, should be considered.

    ii. Limited take-up of the offer of nurse-led follow-up clinics in MOSAICS may have been a barrier to the implementation of optimal care in the intervention practices, but also the pool of patients chosen for both MOSAICS and POST may have been too broad for 'one-size-fits-all' treatment or screening, respectively.

        – One implication is that the outcome of a shift in short-term pain and disability among all OA consulters may be unachievable in primary care, however high quality the care delivered.
        – New ideas, such as identifying patient sub-groups who may selectively benefit from different intensity of primary care treatments, or about different outcome measures (satisfaction with care or improved participation in social life), need to be debated and developed for research and practice with patients who have OA.

Recommendations for future work

This programme has raised issues about the way we think about, treat and measure the experience of patients with OA. Our recommendations are, therefore, grouped into two categories: the need for conceptual thinking and debate, and suggestions for new research avenues.

The conclusions above raise important questions about the need for new conceptual thinking and discussion, recommendations for new research, and wider implementation.

Potential topics for new conceptual thinking and discussion

1. NICE core guidance in practice.

    i. Clarification of the expected benefits of implementing NICE core guidance for patients with OA as a desirable end in itself (e.g. provision of appropriate information and of advice about exercise and pain relief) regardless of whether or not it demonstrably improves clinical outcomes.
    ii. Critical reflection on whether or not current NICE core guidance, even when implemented in the way that MOSAICS has done, results in sufficient shifts in crucial components of patient behaviour to achieve change in clinical outcomes in the short term. An example of this is physical activity for which there is evidence from other research that implementation of physiotherapy-guided

interventions as core treatment could achieve bigger change, so that adding such resource to the MOSAICS package might deliver effects on pain and disability.
iii. Following on from this, review of the delivery of NICE core guidance for primary care and the capacity of a GP- and nurse-led service alone to improve clinical outcomes without additional resources, such as physiotherapy services, to provide individualised and supervised exercise interventions.

2. Continuing debate and critical enquiry about the role, benefits and costs of systematic screening for anxiety and depression in all people with long-term conditions in primary care.
3. The need for new concepts about OA within the research and clinical community.

i. Regarding achievable goals of long-term care for people with this condition, including whether or not it is appropriate to seek more than small short- to medium-term clinical effect sizes in patients' pain and disability in practice. Maintenance of activity and participation despite continuing pain may, for example, be a better long-term measure of effect.
ii. Regarding a combined approach to the management of OA as a long-term condition (importance of good information, adequate advice and resource for exercise and physical activity and other core interventions, identification of individuals at high risk of unfavourable future course, and targeted interventions for those most likely to benefit).

Recommendations for new research

1. Outcomes in long-term conditions.
   Research into new models of long-term care for OA in the context of other long-term disabling conditions that focus on the necessary and desirable process and clinical outcomes from patient, clinician and societal perspectives.
2. Stratified care for OA patients.
   Research to identify subgroups of patients with OA who may, on the basis of combined evidence from previous cohorts, effectiveness studies and the successful implementation strategies described in this programme, benefit from specific interventions in primary care. This includes identification of patient subgroups such as:

   i. those with good prognosis who can be supported to self-manage without additional investigation or treatment
   ii. those who will benefit from specific treatments, such as physiotherapy-led exercise.

3. Exercise and physical activity levels in people with long-term conditions.

   i. Research to identify new approaches to improving long-term adherence to exercise and physical activity among patients with OA such as regular monitoring.

4. Depression and anxiety in people with OA.

   i. Research into more efficient and effective ways of identifying and treating clinically important levels of anxiety and depression in patients with OA.
   ii. Research into the effects of pain management on psychological outcomes in patients with OA.

Wider implementation

Taking the core NICE OA guidance as it stood, we showed it can be implemented in primary care. On that basis, wider implementation with CCGs has begun. However, the hypothesis that it would change pain and disability outcomes was unproven and, in *Chapter 3*, we reflected on one possible explanation for this: that our implementation of NICE guidance in MOSAICS was insufficiently complete or practical. Therefore, the next phase of dissemination incorporates lessons learned in MOSAICS and builds further opportunity

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

for evaluation. We recommend continuing evolution and evaluation of the MOSAICS model of implementation as it is more widely adopted, including addressing and researching the issues raised about the content of the MOSAICS package in *New conceptual thinking and discussion* and *Recommendations for new research*.

Overview of the workstream populations

Workstreams 2–4 investigated different aspects of the primary care management of OA. Although conceptually linked, each was an independent study or set of studies. Each drew on populations based in primary care, but used different practices and different starting points for the selection of the study populations.

In this section, the populations involved in the three workstreams are compared in order to further inform the interpretation and implications of the programme as a whole for all OA patients consulting in primary care by comparing:

1. baseline data across the study samples from the 3 workstreams
2. outcome data across the study samples from the 3 workstreams.

Background

In workstreams 2–4, there were new studies involving a total of 117 general practices and their registered populations. The practices were drawn from the NIHR PCRN in the West Midlands (North) and in the North West.

Comparing baseline characteristics of the recruited populations

Aim

In this section, baseline characteristics of the recruited study populations (irrespective of subsequent allocation arm in the trials) are compared in order to explore (1) selection issues and (2) the nature of OA in primary care populations.

Methods of population selection

To recap and summarise, the selection of each of the samples was as follows:

1. MOSAICS baseline population sample: all persons aged ≥ 45 years, registered with eight participating GPs, were posted a self-complete questionnaire. The response of 53% ($n = 15{,}083$) means that this sample may not be fully representative of all older persons in the general population.
2. MOSAICS trial analysis: all persons aged ≥ 45 years who had responded to the MOSAICS baseline questionnaire, reported pain in at least one joint, consented to follow-up, and who subsequently consulted their GP about joint pain or OA during a 6-month period were included in the trial. This sample of 525 recruited patients represented 26.8% of all patients in these practices consulting about joint pain and OA during this period, and so cannot be assumed to be fully representative of all such patients.
3. The BEEP trial included patients aged ≥ 45 years with knee pain identified in one of three ways: review of medical records in participating general practices to identify persons who had consulted about knee OA in the previous 12 months (69% of participants), referrals (by the GP or self-referral by the patient) to participating physiotherapy services (22.2%), and a screening survey of the registered practice populations for eligible individuals with knee OA (recent consultation not necessary for inclusion) (8.6%). Unlike the MOSAICS trial and POST, the BEEP population was not open to selection by the GP when coding a consultation, but GP selection might operate in terms of which patients they refer for physiotherapy, and patient selection may operate in terms of consent to participate or response to the screening questionnaire. A minimum criterion of pain and disability was applied to all patients for eligibility.

4. The POST included patients aged ≥ 45 years who consulted in 44 GPs about joint pain or OA (an identical set of morbidity or Read codes was used to define inclusion as in the MOSAICS trial). The computer pop-up screen in all practices was fired automatically when the GP entered one of the relevant Read codes and included a set of screening questions (different for intervention and control practices). The GP could bypass these questions and selectivity in trial recruitment might occur. However, unlike the MOSAICS trial, there was no selection prior to consultation by a pre-trial survey.

Therefore, each study population had possibilities for selectivity of the recruited versus their target population of patients with clinical OA, and some of these differed between the studies. However, there were also logical reasons why the recruited populations might differ between the studies:

1. Persons consulting with their GP about joint pain or OA compared with persons in the general population responding in a questionnaire that they have joint pain or OA compared with patients referred (by GP or by self) to physiotherapy because of OA.
2. Persons consulting with their GP about joint pain and OA who had previously consented to follow-up compared with persons consulting their GP about joint pain and OA with no prior selection by consent or survey response compared with persons with knee pain consenting to a physiotherapy-led intervention.

Results

Table 59 compares some baseline characteristics in the recruited populations regardless of the arm to which they (or their practices) were allocated in the trials. We have used subgroups with knee pain for some of the comparisons in order to include all the main studies from the three workstreams. Such subgroups include most people in the relevant study sample with joint pain or OA, although they constituted only about half of the POST sample.

Discussion

Despite the variation in sampling strategies for the different studies, the populations of people aged ≥ 45 years who have joint pain and/or diagnosed OA do not differ dramatically in sociodemographic characteristics, except that the MOSAICS trial population had a higher proportion of people out of work than all the other study populations, including the wider population from which it was drawn (i.e. the MOSAICS population survey). This means that the differences cannot arise simply from demographic contrasts between the general practices selected for the different workstreams.

Specific similarities and differences

1. General population samples: persons with knee pain compared with persons without knee pain based on self-reported questionnaire (the MOSAICS population survey).
   As would be expected from published population surveys (e.g. O'Reilly *et al.*[349] and Jinks *et al.*[350]), persons > 45 years reporting knee pain are a little older, are more likely to be female, have a higher mean BMI, are less likely to be employed, are more likely to have anxiety and depression, and are more likely to have lower EQ-5D health status than persons reporting that they have not experienced knee pain recently. The data have not been adjusted for age and sex, so the differences in the other variables might be explained by age and sex.
2. Persons reporting joint pain in the general population (MOSAICS population survey) compared with persons recruited when consulting their GP with joint pain or clinical OA (MOSAICS trial).
   Consulters with knee pain recruited into the MOSAICS trial were, with the exception of one variable, different from persons who reported knee pain in the baseline MOSAICS practice population survey: they were a little older, more likely to be female, more likely to have a higher BMI, less likely to be employed, and more likely to have a lower EQ-5D health status. This would be consistent with consulters representing a subgroup with generally more severe problems than all persons reporting knee pain in the general population. However, the possibility that GPs selectively recruited consulting patients into the trial who had rather more severe problems cannot be ruled out.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 59 Baseline comparability of study populations

	Study name					
	BEEP	POST	MOSAICS population survey			MOSAICS patient-level trial
Study sample size	(*N* = 514)[a]	(*N* = 1412)[a]	All (*N* = 15083)[b]	Knee pain (*N* = 8159)[a,c]	No knee pain (*N* = 6751)[b,c]	(*N* = 525)[a]
Age (years), mean (SD)	63.0 (10.0)	65.5 (10.3)	63.4 (11.2)	64.3 (11.2)	63.3 (11.2)	67.3 (10.5)
Sex, *n* (%)						
Female	262 (51.0)	799 (56.6)	8198 (54.4)	4551 (55.8)	3551 (52.6)	313 (59.6)
Male	252 (49.0)	613 (43.4)	6885 (45.6)	3608 (44.2)	3200 (47.4)	212 (40.4)
Marital status, *n* (%)						
Married	376 (73.9)	974 (69.1)	10345 (69.0)	5547 (68.4)	4693 (69.8)	354 (67.7)
Separated	6 (1.2)	20 (1.4)	203 (1.4)	112 (1.4)	90 (1.3)	6 (1.1)
Divorced	45 (10.6)	115 (8.2)	1114 (7.4)	624 (7.7)	476 (7.1)	42 (8.0)
Widowed	45 (10.6)	194 (13.8)	1879 (12.5)	1087 (13.4)	757 (11.3)	81 (15.5)
Cohabiting	19 (4.5)	51 (3.6)	661 (4.4)	336 (4.1)	319 (4.7)	19 (3.6)
Single	18 (4.5)	56 (4.0)	791 (5.3)	398 (4.9)	386 (5.7)	21 (4.0)
BMI (kg/m^2), mean (SD)	29.6 (5.7)	28.7 (5.4)	26.9 (4.7)	27.6 (5.0)	26.1 (4.2)	28.3 (5.0)
Employed, *n* (%)						
Yes	214 (42.2)	452 (32.0)	5330 (36.1)	2650 (33.3)	2639 (39.8)	136 (26.3)
No	293 (57.8)	959 (68.0)	9429 (63.9)	5317 (66.7)	3993 (60.2)	381 (73.7)
Knee pain, *n* (%)	514 (100)	778 (55.1)	8159 (54.7)	–	–	433 (82.5)
Knee pain score (NRS),[d] mean (SD)	6.8 (1.7)	6.6 (2.1)	–	4.4 (2.6)	–	6.6 (2.4)
PHQ-8, mean (SD) (depression)	4.1 (4.8)	6.1 (6.1)	–	–	–	4.8 (5.0)
GAD-7, mean (SD) (anxiety)	3.4 (4.6)	5.3 (5.7)	3.5 (4.7)	4.7 (5.3)	3.0 (4.4)	3.5 (4.7)
EQ5D, mean (SD) (general health status)	0.64 (0.23)	0.58 (0.24)	0.75 (0.26)	0.68 (0.28)	0.84 (0.20)	0.58 (0.28)

It should be noted that the figures in the table may differ from those given in the individual chapters where the latter are presented as median and IQRs rather than, as here, means and SDs.

a Populations complaining of joint pain or diagnosed with OA.

b Mixed populations including people with and without joint pain or OA.

c 173 responders did not answer the question on knee pain.

d Knee pain score among participants with knee pain at consultation. Pain score refers to average in last 6 months for BEEP and in last 3 months for MOSAICS and POST. For BEEP, pain severity score was only available for participants recruited via the medical record review or postal survey methods (*n* = 409).

The exception to this general pattern was that anxiety levels were lower in the consulters than in persons with knee pain in the general population; this is consistent with previous findings that suggest that psychological distress is associated with lower rates of consultation among knee pain patients and may act as a barrier to consultation.[351]

3. Persons recruited when consulting their GP with joint pain or clinical OA (MOSAICS trial) compared with patients with knee pain recruited from primary care into a physiotherapy trial (BEEP trial).

The BEEP trial participants were younger, less likely to be female, more likely to be employed and more likely to have lower depression scores, had similar anxiety scores to MOSAICS consulters and had better EQ-5D health status than MOSAICS consulters. All of this suggests that BEEP trial participants represented a selectively healthier or rather less severe group than OA consulters in general. However, the mean knee pain severity score of BEEP trial participants at baseline was similar to that of MOSAICS consulters with knee pain. This is compatible with the recruitment criteria for BEEP (a minimum score on the Chronic Pain Grade). The better general health and higher BMI of BEEP participants may reflect selection factors for referral to physiotherapy and recruitment into physiotherapy trials generally. Some of the differences may also represent demographic contrasts between the registered populations of the practices recruited by the two workstreams, for example the proportions of people currently in employment.

4. All persons consulting their GP with joint pain or clinical OA who had a screening template fired (POST) compared with persons recruited when consulting their GP with joint pain or clinical OA who had consented in the earlier population survey (MOSAICS trial) compared with all persons reporting knee pain in the population survey (MOSAICS population survey knee pain subgroup).
   For some characteristics (namely mean age, proportion of females, proportion of married persons, proportion of employed persons), POST participants' values were closer to those of the MOSAICS population survey knee pain sample than were the values of the MOSAICS trial sample, suggesting that they were more generally representative of older people with troublesome joint pain in the general population. However, mean BMI score, mean EQ-5D health status and mean knee pain severity score among those with knee pain were closer in the POST sample to the means for these variables in the MOSAICS trial participants than to knee pain sufferers in the general population survey. This is consistent with consulters generally representing a subgroup of everyone in the older population with currently more severe pain and poorer general health.
   The one major contrast between POST participants and all other study groups was that this population had higher psychological distress on both depression and anxiety scales. In terms of criteria for inclusion in the cluster randomised POST, the expectation was that mean anxiety and depression might be lower in POST participants than in all other study groups, as patients with severe mental health problems were excluded and all other consulters with joint pain and OA should have been screened. The observation that distress was higher in POST participants than in all other study groups raises the possibility that the purpose of the cluster trial (screening in the intervention group for anxiety and depression) may have influenced GP behaviour (in the intervention group at least) to selectively include patients with more severe problems. Such selectivity has been highlighted before as a potential source of bias in cluster RCTs.[187]

Conclusion

The purpose of this section was to present comparative baseline data from primary workstream studies as the basis for discussion, given that all these study populations were considered to represent primary care patients with OA in one way or another. Each study had different objectives and different methods of sampling and it was likely that different populations would be recruited.

The size of the crude differences are not large, which is important given the number of practices involved and the many potential opportunities for variation between practices and for different types of selectivity to be operating. As *Table 59* makes clear, many of the differences are consistent with a simple explanatory framework that persons aged ≥ 45 years who consult primary health-care services about joint pain are likely to experience more severe symptoms and a greater impact on their lives at the time of consultation than all persons who report current joint pain at any one time in the general population.

This explanatory framework should not have implications for the internal validity of the individual workstream studies, but does raise critical questions about the potential of intervention trials in patients with long-term conditions characterised by intermittent exacerbations that recruit at, or close to, the time of consultation to demonstrate a post-consultation effect when there is likely to be post-consultation improvement in mean pain and disability in all consulters.[352,353]

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

The discussion above has also considered how some of these differences may have arisen because of conscious or unconscious recruitment selectivity by clinicians of more severely affected patients or by patients invited into exercise and activity trials. This should not have implications for the internal validity of the BEEP trial, but needs to be considered in relation to the internal validity of the cluster trials.

Response to care

In one cluster trial (MOSAICS) and in one individually randomised trial (BEEP), clinical outcomes of pain and disability were not improved in patients aged ≥ 45 years with clinical OA by new interventions. However, all patients recruited were, in all arms of both trials, receiving components of OA care that might not either be always routinely accessed (BEEP) or be available (MOSAICS) in current daily practice outside the arena of the trials, and which were separate to the novel interventions being tested in the studies. These included use of an e-template to record quality of care in consultations affecting implementation of NICE guidance and physician behaviour (MOSAICS), and physiotherapy-led exercise and physical activity for knee OA (BEEP). We looked at the response to care of exercise and physical activity across these two trials.

Aim

To compare 'response to care' across two trials, which included advice or action to support change in exercise and physical activity behaviour among patients with OA as a component of all arms.

Method

The PASE scores and the change in these scores between 0 and 6 months were summarised and compared between the two trials for all participants in intervention and control groups.

Results

Table 60 shows relevant variables at 0 and 6 months in the two studies.

Discussion

The PASE scores were different between the two trials, with patients seeing the physiotherapist in BEEP scoring higher at the start of the trial than MOSAICS patients immediately after their index consultation. However, there is also a clear difference between MOSAICS and BEEP trial participants in the extent of change from baseline to 6 months of mean PASE scores. This is regardless of intervention or control status within the two studies. There were increases in mean PASE scores of similar extent in all BEEP trial arms, contrasted with declines in mean PASE scores among MOSAICS participants. This was explored in the MOSAICS chapter, with walking identified as the component that declined in the intervention group while muscle-strengthening exercises were reported as improving. Despite the lack of additional effect of the new physiotherapy interventions, these overall findings suggest that physiotherapy-led exercise and activity could enhance the NICE core guidance delivered by the MOSAICS model of care.

TABLE 60 Response to care in three trial populations: mean PASE scores and their change over time

		Start of the trial		6-month follow-up		Difference 6 months minus 0 months	
Variable	Group	MOSAICS	BEEP	MOSAICS	BEEP	MOSAICS	BEEP
PASE	Control	147.49	176	136.24	188	–1.25	+ 12
	Intervention (1)	138.71	175	123.02	189	–5.69	+ 14
	Intervention (2)		180		196		+ 16

Higher PASE score denotes greater levels of physical activity.

Outputs from delivery and dissemination of the programme

The programme included detailed study development, training programmes, project management and the delivery of three trials, recruiting 2454 patients. The studies within the programme were supported by 117 general practices (members of the West Midlands North and Cheshire and Merseyside Comprehensive CLRNs) and all five physiotherapy services in Staffordshire, Shropshire and South Cheshire. In this section, we describe the partnerships that supported this work, innovations that were developed and applied to support successful delivery of the programme, and key dissemination activities linked to, and arising from, the programme. The PPI that has underpinned all phases of the programmes has been described elsewhere in the report (see *Appendix 1* and in the individual workstream chapters: see *Chapters 3–5*).

Partnerships and collaborative appointments

The North Staffordshire Primary Care Research Consortium

This is the body that formalises the research partnership in place between Keele and its NHS partners, including North Staffordshire Commissioning Group (CCG), Stoke CCG, South and West Cheshire CCGs, lead research active service GPs and physiotherapists, and public health and the community provider unit in Staffordshire. The Consortium's Board oversees joint University–NHS research activities, with the aim of ensuring that the strategy reflects NHS and patient priorities and needs, and contributes to building research capacity in the NHS and active NHS participation in delivering the research. North Staffordshire CCG is the lead organisation in the Consortium and held the funding contract for this NIHR programme.

The Primary Care Research Network and the Primary Care delivery arms of West Midlands North and Cheshire and Merseyside Comprehensive Research Networks

The NIHR CLRNs provided crucial infrastructure support and funding to enable NHS patients, GPs, research nurses, practice nurses, physiotherapists and NHS managers to participate in, and benefit from, research. The aim of the NIHR CLRN is to make the NHS a world-class environment in which to undertake research by promoting a culture of research as a normal part of everyday clinical practice. Through the resources provided by the CLRNs and innovative joint working between the Networks, Keele University and the Consortium, a GP Research Network of 117 GP practices and 185 individual GPs, and a Physiotherapy Research Network comprising all five physiotherapy services in our region and 74 individual physiotherapists, were involved in delivering this large NIHR programme. Effective joint working across the NHS/network/university interface resulted in development of innovative approaches to support fast and effective recruitment of patients, early identification of barriers and solutions to overcome them, and early adoption of trial findings to improve clinical practice.

New GP research facilitator ($n = 5$ posts) and physiotherapy research facilitator posts ($n = 5$) were established as joint academic/network/NHS appointments, who helped to:

- disseminate the innovative methods we developed to support busy NHS clinicians to recruit patients to the research
- contribute to developing the intervention packages that were tested
- identify and recruit GP and physiotherapy centres
- train NHS clinicians in the research procedures
- provide ongoing support for the delivery of specific projects.

Challenges and innovation in research delivery

Example: workstream 3 – the BEEP trial

This trial was based on extensive engagement of NHS staff and organisations across the North West Midlands and South Cheshire, enabling successful recruitment ahead of target of 526 patients from 65 GP practices and five physiotherapy services, involving clinics run by 47 physiotherapists.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Innovation in methods of research delivery

Crucial to this engagement was the combination of infrastructure support provided by the CLRNs, the creation of joint academic service physiotherapy research facilitator posts appointed between Keele, the CLRN and the NHS, and the quality of engagement in the trial by local physiotherapy leads and clinical champions (which was as a result of the support they received from the facilitators and the North Staffordshire Primary Care Research Consortium).

Recruitment to the trial was highlighted as a challenge in the pilot and was successfully addressed in the main trial through carefully planned coworking between NHS, Network and university staff. We established new triage procedures in the physiotherapy referral pathway to identify those patients who were potentially eligible to take part and direct them to research clinics. We also undertook regular searches of GP consultation systems and clinical screening of those patients who had consulted with knee pain, to ensure delivery of the right patients to the right physiotherapists. We provided support for NHS staff to be released to receive training and periodic updates on the research interventions from the Keele team and the research facilitators provided hands-on support and mentoring for all sites. As a result, this large trial screened 5345 patients and recruited 514 people to the trial 2 months ahead of target.

A study dissemination day was held at the end of the trial for all physiotherapists and general practices involved in the study.

Challenges

One key challenge was posed by the disappointment of intervention physiotherapists when they were informed of the negative trial finding (i.e. that their enhanced exercise training schedules had not, after all their effort, resulted in better long-term adherence to exercise or in better clinical outcomes than those achieved by the group with physiotherapists delivering usual-care exercise and physical activity advice). There was a clear feeling that an enhanced package had been delivered that made sense to the physiotherapists. However, there was also a clear sense that the 'usual-care' physiotherapy had been delivered well and, under the auspices of the trial, in optimal circumstances with optimal support (compared with the variation observed in the usual daily NHS context). The practical challenge now is how to improve care pathways so that the right patients benefit from the right kind of physiotherapy intervention in everyday practice. Although physiotherapy exercise and activity advice has a strong evidence base for effectiveness, the research challenge that remains is how to improve and maintain changes in exercise and physical activity behaviour sufficiently to influence the long-term course of OA.

Example: workstream 4 – POST

The practical challenge here was different. We adopted a light touch approach to training clinicians and provided practice-level e-templates for GPs to screen patients in real-time consultations as the main intervention, with a pragmatic approach to the resulting treatment.

Innovation in methods of research delivery

A total of 44 practices were recruited and 150 GPs were supported to undertake systematic screening of patients consulting with OA or joint pain over a 4-month period. 'Pop-up' electronic templates were installed on GPs' computerised consultation systems, which automatically fired on entry of Read codes associated with patient consultation for OA or joint pain conditions. The templates prompted GPs to check patients' eligibility to be included in the trial, to screen for pain intensity (control practices) and for anxiety and depression (intervention practices). Once completed, the templates added a trial recruitment code to the patient record (to prevent duplication of recruitment and data collection). If the patient was deemed ineligible, the template provided a prompt for the GP to record reason for exclusion from the trial.

The trial recruited 1415 patients on time. Trial recruitment and data collection procedures were embedded into real-time GP consultation systems, reducing the administrative burden on GPs when completing research and enabling the research team to undertake real-time audit and monitoring of recruitment and research data collection.

Dissemination

Example: workstream 2 – the MOSAICS suite of studies

This workstream identified and evaluated methods to improve and support implementation of NICE OA guidance in primary care and to improve the quality of primary care for patients with OA. On the basis of this evidence, dissemination activities have followed.

Study feedback day (autumn 2013) and associated regional and national quality, training and implementation initiatives

After completion of the follow-up and analysis of the MOSAICS workstream described in *Chapter 3*, a study feedback day was held for CCG representatives and all the health-care professionals, research support professionals, RUG members and research team members who had been involved in the MOSAICS workstream studies within the eight participating general practices. Separate generic feedback was provided to participating patients through newsletters and poster displays in the practices.

A major outcome from the study day was that the GP clinical lead for Shropshire CCG led action to ensure that the methods of implementing evidence that had formed our research was incorporated into quality initiatives undertaken in a locality group of practices within the CCG. This included:

- The South Shropshire locality group (14 general practices) partnering with Keele to submit a successful application to the NHS Regional Innovation Fund to adapt the MOSAICS intervention for roll-out across the locality group (the JIGSAW initiative). The CCG supported the appointment of three GP clinical champions and two practice nurse-led clinical champions, which ensured buy-in across the 14 practices. Keele provided training and mentoring support for clinicians and nurses on the MOSAICS intervention, the patient guidebook and the e-template to guide quality care for OA. The MOSAICS training was adapted to a more practical, 2-day duration for practice nurses and 1-day duration for GPs. The e-template can be used in addition to monitor and audit subsequent primary care clinical management of patients with OA and to facilitate continuing improvements in practice.
- This involved or linked to other dissemination initiatives to improve current clinical practice associated with the NIHR programme and personnel.
    - All participating GPs from the locality group have undertaken the OA section of the Royal College of General Practitioners' online musculoskeletal module, a national initiative developed by collaboration between Arthritis Research UK and Keele, which includes components from the MOSAICS training programmes developed by Porcheret *et al.*[134,143]
    - Nurse training has been conducted by a team from the NIHR programme under a regional initiative of the national body for training in the NHS (Education for Health, Warwick, UK) focused on nursing.
    - Adoption by the West Midlands AHSN's innovation programme for long-term conditions, as part of their aim of adopting innovation to improve clinical practice. The AHSN has expressed an interest in:
        - rolling out the MOSAICS implementation model to a second locality group (22 general practices in Telford and Wrekin)
        - supporting an initiative for non-surgical alternatives to arthroscopy led by Keele's NICE Fellow (Professor Krysia Dziedzic, principal investigator of the MOSAICS workstream) and the University Hospital of Coventry and Warwick's NICE Scholar (Mr Tim Barlow, Orthopaedic Surgeon)
        - further work on developing and measuring the effectiveness of the pathway of care for OA involving physiotherapy and general practice led by the North Staffordshire Primary Care Research Consortium's manager (Ms Helen Duffy).

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

National and international dissemination

1. Conference presentations.

    i. Chartered Society of Physiotherapists conference 2014: programme personnel and OA RUG representative presented the MOSAICS workstream and the JIGSAW initiative.
    ii. Royal College of General Practitioners' annual conference 2012: the training programme for the GPs was presented in a workshop session to an audience of 200 GPs.
    iii. The annual named Droitwich lecture given at the British Society for Health Professionals in Rheumatology was given in 2014 by Professor Dziedzic, principal investigator of the MOSAICS workstream, and was largely based on the MOSAICS work.

2. Training packages led or written by programme personnel drawing on content of the training developed in MOSAICS.

    i. Royal College of General Practitioners' online musculoskeletal training for primary care: OA module (see also *Study feedback day (autumn 2013) and associated regional and national quality, training and implementation initiatives*)
    ii. clinical knowledge summary on OA for clinicians: NICE-linked evidence-based resource
    iii. *British Medical Journal* educational module on OA for GPs
    iv. EULAR teaching module on OA for rheumatologists.

3. Resources.

    i. The NIHR programme is informing the work of NICE in introducing quality standards for its OA guideline in the absence of a quality of care framework for OA.

Challenges

The MOSAICS programme of development and implementation of training, quality initiatives, patient information resource and quality of care monitoring for the primary care management of the patient with OA has struck a chord with professional leaders and educators, patients and some clinicians. We have provided evidence that the MOSAICS programme delivers better information for patients, better systems for improving and monitoring the quality of care delivered in the GP consultation, and improved use of simple analgesia, physiotherapy referral and radiography. These are components of the evidence-based care of long-term symptoms and diseases, and their improved quality and delivery by MOSAICS is therefore a desirable outcome in its own right. We have completed successful dissemination of these components and achieved high take-up in the NHS and with patients because these approaches do not lead to excessive new costs. However, the challenge remains to understand why such activity is not appearing to improve the desirable clinical outcomes of reducing pain and improving physical function. There remain practical barriers to overcome in implementing or adapting features of implementation (such as the e-template), which may be better delivered in the context of a short-term supervised trial than in the long term of real-life practice.

Acknowledgements

Contributions of authors

(Unless otherwise stated, all job titles relate to the Research IPCHS at Keele University.)

Professor Elaine Hay (Professor of Community Rheumatology and Director of the Arthritis Research UK Primary Care Centre, Chief Investigator of the whole programme) was responsible for the conception, design, interpretation, drafting and revising of the whole programme.

Professor Krysia Dziedzic (Arthritis Research UK Professor of Musculoskeletal Therapies and NICE Fellow) was the Principal Investigator of the MOSAICS studies (see *Chapter 3*) including conception, design, interpretation, drafting and revision.

Professor Nadine Foster (NIHR Professor of Musculoskeletal Health in Primary Care) was the Principal Investigator of the BEEP studies (see *Chapter 4*) including conception, design, interpretation, drafting and revision.

Professor George Peat (Professor of Clinical Epidemiology) was the Principal Investigator of the POST study (see *Chapter 5*) including conception, design, interpretation, drafting and revision.

Professor Danielle van der Windt (Professor of Primary Care Epidemiology) wa the Principal Investigator of workstream 1 studies (see *Chapter 2*) including conception, design, interpretation, drafting and revision.

Dr Bernadette Bartlam (Lecturer in Health Services Research) was responsible for the data acquisition, analysis, interpretation, drafting and revising of qualitative studies within POST (see *Chapter 5*).

Dr Milisa Blagojevic-Bucknall (Lecturer in Statistics) was responsible for the analysis, interpretation, drafting and revising of workstream 1 studies (see *Chapter 2*).

Dr John Edwards (GP Research Fellow) was responsible for the design, data acquisition, interpretation, drafting and revising of the population-level GP database studies in MOSAICS (see *Chapter 3*).

Dr Emma Healey [Post-doctoral Research Fellow (Clinical and Health Services Research)] was responsible for the data acquisition, interpretation, drafting and revising of component studies in MOSAICS and BEEP (see *Chapters 3* and *4*).

Dr Melanie Holden (Arthritis Research UK Allied Health Professional Training Fellow) was responsible for the design, data acquisition, interpretation, drafting and revising of BEEP studies (see *Chapter 4*).

Ms Rhian Hughes [Co-director of the Research Institute for Primary Care and Health Sciences (Research and Health Services Co-ordination and Support)] was responsible for the conception, design and lead on patient recruitment and data acquisition for the whole programme. Was also responsible for the drafting and revising of *Chapter 6*.

Dr Clare Jinks [Senior Lecturer in Health Services Research (Social Science and PPI)] was responsible for the conception, design, analysis, interpretation, drafting and revising of qualitative studies within BEEP (see *Chapter 4*) and of the PPI component of the whole programme (see *Appendix 1*).

Professor Kelvin Jordan (Professor of Biostatistics) was responsible for the design, analysis, interpretation, drafting and revising of the population-level GP database studies in MOSAICS (see *Chapter 3*).

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Dr Sue Jowett (Senior Lecturer in Health Economics and Health Economic Modelling at the Health Economics Unit, University of Birmingham and Honorary Research Fellow at Keele University) was the health economics lead for the programme, including design, analysis, interpretation, drafting and revising of the health economic components of workstream 1, MOSAICS, BEEP and POST (see *Chapters 2–5*).

Dr Martyn Lewis (Reader in Biostatistics) was responsible for the design, analysis, interpretation, drafting and revising of the individual-level MOSAICS cluster trial and POST (see *Chapters 3* and *5*).

Professor Christian Mallen (NIHR Professor of General Practice Research) was responsible for the clinical research lead for POST including conception, design, interpretation, drafting and revising (see *Chapter 5*).

Dr Andrew Morden (Research Associate in Health Services Research) was responsible for the design, data acquisition, analysis, interpretation, drafting and revising of qualitative studies in MOSAICS (see *Chapter 3*).

Ms Elaine Nicholls (Research Assistant in Biostatistics) was responsible for the analysis, interpretation, drafting and revising of BEEP studies (see *Chapter 4*).

Professor Bie Nio Ong [Emeritus Professor of Health Services Research (Social Science and PPI)] was the social science lead for the whole programme, including PPI (see *Appendix 1*), and was responsible for the and conception, design, data acquisition, analysis, interpretation, drafting and revising of the MOSAICS studies (see *Chapter 3*).

Dr Mark Porcheret (Senior Lecturer in General Practice) was responsible for the conception, design, data acquisition, analysis, interpretation, drafting and revising of the GP training and evaluation programme of MOSAICS (see *Chapter 3*).

Mr Jerome Wulff (PhD Student in Health Economics) was responsible for using his PhD work to form the basis for workstream 1 (see *Chapter 2*), including data preparation, analysis, interpretation and drafting.

Dr Jesse Kigozi (Research Fellow, Health Economics Unit, University of Birmingham) carried out supervised data organisation, analysis and interpretation for the health economic studies in the BEEP trial.

Dr Raymond Oppong (Research Fellow, Health Economics Unit, University of Birmingham) carried out supervised data organisation, analysis and interpretation for the health economic studies in MOSAICS.

Dr Zoe Paskins (Clinical Lecturer and Honorary Consultant in Rheumatology) was responsible for the GP training, qualitative work in MOSAICS and also developed and completed a PhD linked to the programme.

Professor Peter Croft (Professor of Primary Care Epidemiology) contracted grant holder and original chief investigator for the programme and was responsible for the conception, design and interpretation of programme plus drafting of the overall summary, introduction and sections of *Chapter 6*.

Trial Steering Committee and Data Monitoring Committee members

MOSAICS Trial Steering Committee Membership

Professor Philip Conaghan (chairperson), Professor of Musculoskeletal Medicine, Leeds University, Leeds, UK.

Professor Anne Rogers, Professor of Health Systems Implementation, University of Southampton, Southampton, UK.

Professor Sandra Eldridge, Professor of Biostatistics, Barts and The London School of Medicine and Dentistry, London, UK.

Dr Marta Buszewicz, Reader in Primary Care, University College London, London, UK.

Professor Alison Hammond, Professor in Rheumatology Rehabilitation, University of Salford, Salford, UK.

Dr Sandra Hollinghurst, Senior Lecturer in Health Economics, University of Bristol, Bristol, UK.

Ms Jo Cumming, Helplines Manager, Arthritis Care, London.

Ms Cynthia Mellish, Registered General Nurse, Practice Nurse Independent Nurse Prescriber, Stockport PCT, Stockport, UK.

Professor Elaine Hay, Professor of Community Rheumatology, Keele University, Keele, UK.

Patient and public involvement (research user group) members

Carol Ingram.

Christine Walker.

POST Trial Steering Committee membership

Dr Caroline Mitchell (chairperson), Senior Clinical Lecturer, University of Sheffield, Sheffield, UK.

Dr Sally Kerry, Reader in Medical Statistics, Barts and the London Medical School, London, UK.

Dr Darren Carr, Consultant Psychiatrist and Clinical Director for the Neuropsychiatry and Old Age Psychiatry Directorate (Moorlands), North Staffordshire Combined Healthcare NHS Trust, Stoke-on-Trent.

Patient and public involvement (research user group) members

Chris Pope.

Josephine Bird.

BEEP Trial Steering Committee Membership

Professor Mike Hurley (chairperson), Professor of Rehabilitation Sciences, Kingston University and St George's, London, UK.

Dr Christopher Gidlow, Associate Professor, Centre for Sport, Health and Exercise Research, Staffordshire University, Staffordshire, UK.

Dr Adrian Pendleton, Consultant Rheumatologist, the Royal Victoria Hospital and Musgrave Park Hospital, Belfast.

Patient and public involvement (research user group) members

John Murphy.

Bernard Colcough.

Mavis Dutton.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Data Monitoring Committee membership

BEEP (Data Monitoring Committee 1)

Professor Chris Roberts (chairperson), Professor of Biostatistics, University of Manchester, Manchester, UK.

Dr Christina Jerosch-Herold, Reader, School of Health Sciences, University of East Anglia, Norwich, UK.

Professor James Selfe, Professor of Physiotherapy, University of Central Lancashire, Preston, UK.

Professor Richard McManus, Professor of Primary Care, University of Oxford, Oxford, UK.

MOSAICS and POST (Data Monitoring Committee 2)

Dr Janine Grey (chairperson), Principal Statistician, Clinical Trials Research Unit, University of Leeds, Leeds, UK.

Dr Jane Daniels, Deputy Director, CTU, University of Birmingham, Birmingham, UK.

Dr Mike Callaghan, Research Fellow/Clinical Specialist Physiotherapist, Arthritis Research UK Epidemiology Unit, University of Manchester, Manchester, UK.

Professor Walter Gregory, Professor of Statistical Methodology in Clinical Trials, University of Leeds, Leeds, UK.

Programme personnel

Programme chief investigator

Elaine Hay (2011–14).

Principal investigators

Krysia Dziedzic.

Nadine Foster.

Christian Mallen.

George Peat.

Danielle van der Windt.

Other programme coapplicants

Josephine Bird.

Stirling Bryan. (Stirling Bryan left Birmingham to take up a post in Canada in 2010. The Birmingham Health Economics Unit continued to provide collaboration and support with the team for this programme led by Sue Jowett.)

Peter Croft (programme chief investigator 2009–10).

Helen Duffy.

Rhian Hughes.

Bie Nio Ong.

Elaine Thomas.

Programme management

Elizabeth Mason.

Patient involvement

Bernard Colclough.

Barbara Lambert.

Alan Sutton.

The late Brian Dudley.**

Gordon Lambert.

Robert Taylor.

Mavis Dutton.

John Murphy.

Susan Vaughan.

Kathy Fell.

Chris Pope.

Anne Worrall.

Teresa George.

Carol Rhodes.

Adele Higginbottom.

Jeanette Shipley.

**Brian Dudley passed away just before the programme was completed.

Researchers

Ebenezer Afolabi.

Sue Hill.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Bie Nio Ong.

Bernadette Bartlam.

Clare Jinks.

Jon Packham.

John Belcher.

Kelvin Jordan.

Zoe Paskins.*

Lauren Brooks.

Sue Jowett.

Mark Porcheret.*

John Buckley.

Jesse Kigozi.

Diane Roberts.

Milisa Blagojevic-Bucknall.

Martyn Lewis.

Ed Roddy.

Rebecca Case.

Chris J Main.

Sarah Ryan.

Mike Doherty.

Andrew Moore.

Robert Smith.*

John Edwards.*

Ricky Mullis.

Gail Sowden.

Andrew Finney.*

Gretl McHugh.

Elaine Thomas.

Daniel Green.*

Andrew Morden.*

David Whitehurst.

Emma Healey.

Barbara Nicholl.

Jerome Wulff.*

Mel Holden.

Elaine Nicholls.

*Undertook or are undertaking PhDs within or linked to the programme, in addition to or including their work on the programme.

Medical education advisory group

Andrew Hassell.

Robert McKinley.

Study co-ordinators

Kris Clarkson.

June Handy.

Barbara Nicholl.

Angela Pushpa-Rajah.

Stephanie Tooth.

Julie Young.

Administration and informatics

Claire Ashmore.

Jo Jordan.

Jenny Stevens.

Jo Bailey.

Zoe Lingard.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Ian Thomas.

Natalie Barcroft.

Liz Hartshorne.

Sally Thomas.

Alicia Bratt.

Charlotte Purcell.

Simon Wathall.

Clare Calverley.

Tracy Reynolds.

Tracy Whitehurst.

Lucy Gibbons.

Michelle Robinson.

Nesta Williams.

Sue Hunt.

Primary Care Research Network and practice-based staff

Jessica Graysmark.

Marvin Ryles.

Sally Thomas.

Research nurses

Deborah D'Cruz.

Kathryn Dwyer.

Will Meredith.

Rebecca Rider.

Kanchan Vohora.

Julie Young.

Clinical research facilitators (physiotherapy and general practitioner)

Carol Doyle.

Tina Hadley-Barrows.

Treena Larkin.

Vije Rajput.

Yvonne Rimmer.

Physiotherapy managers, BEEP clinical leads and trial physiotherapists

Ross Anderson.

Kate Harper.

Heather Scholes.

Alastair Baker.

Ruth Heaton.

Nishu Sehijpal.

Davelee Brocklehurst.

Carla Horricks.

Russell Shill.

Steve Burton.

Philip Hulse.

Stephen Siveter.

Stuart Cameron.

Katrina Humphreys.

Chris Smith.

Dave Clark.

Beth Keddie.

Nicola Stephenson.

Libby Cockill.

Zena Kelly.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Sheila Stringer.

Tandi Coetzee.

Andy Laing.

Liz Stuart.

Esther Collacott.

Lynn Martin.

Helen Thompson.

Sue Cross.

Julie McLauchlan.

Sam Townsend.

Jason Curtis.

Alexandra Needham.

Steph Turner.

Caroline Dady.

Mandy Onyet.

Julie Walker.

Mike Delahay.

Dean Peer.

Mike Watton.

Eddie Dovison.

Sue Perks.

Lone Odgaard Webber.

Michelle Evans.

Claire Powell.

Lis Wharton.

Joanne Freelove.

Linda Pugh.

Greg Wilary.

Hazel Garside.

Rowena Quirk.

Laura Williams.

Stephanie Gommersall.

Helen Saunders.

Noel Harding.

Stuart Scattergood.

Publications

Jordan JL, Holden MA, Mason EE, Foster NE. Interventions to improve adherence to exercise for chronic musculoskeletal pain in adults. *Cochrane Database Syst Rev* 2010;**1**:CD005956.

Morden A, Jinks C, Ong BN. Lay models of self-management: how do people manage knee osteoarthritis in context? *Chronic Illn* 2011;**7**:185–200.

Dziedzic KS, Healey EL, Main CJ. Implementing the NICE osteoarthritis guidelines in primary care: a role for practice nurses. *Musculoskeletal Care* 2013;**11**:1–2.

Finney A, Porcheret M, Grime J, Jordan KP, Handy J, Healey E, *et al.* Defining the content of an opportunistic osteoarthritis consultation with primary health care professionals: a Delphi consensus study. *Arthritis Care Res* 2013;**65**:962–8.

Porcheret M, Grime J, Main C, Dziedzic K. Developing a model osteoarthritis consultation: a Delphi consensus exercise. *BMC Musculoskelet Disord* 2013;**14**:25.

Uthman OA, van der Windt DA, Jordan JL, Dziedzic KS, Healey EL, Peat GM, Foster NE. Exercise for lower limb osteoarthritis: systematic review incorporating trial sequential analysis and network meta-analysis. *BMJ* 2013;**347**:f5555.

Dziedzic KS, Healey EL, Porcheret M, Ong BN, Main CJ, Jordan KP, *et al.* Implementing the NICE osteoarthritis guidelines: a mixed methods study and cluster randomised trial of a model osteoarthritis consultation in primary care – the Management of OsteoArthritis In Consultations (MOSAICS) study protocol. *Implement Sci* 2014;**9**:95.

Foster NE, Healey EL, Holden MA, Nicholls E, Whitehurst DG, Jowett S, *et al.* A multicentre, pragmatic, parallel group, randomised controlled trial to compare the clinical and cost-effectiveness of three physiotherapy-led exercise interventions for knee osteoarthritis in older adults: the BEEP trial protocol (ISRCTN: 93634563). *BMC Musculoskelet Disord* 2014;**15**:254.

Morden A, Jinks C, Ong BN, Porcheret M, Dziedzic KS. Acceptability of a 'guidebook' for the management of Osteoarthritis: a qualitative study of patient and clinician's perspectives. *BMC Musculoskelet Disord* 2014;**15**:427.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Ong BN, Morden A, Brooks L, Porcheret M, Edwards JJ, Sanders T, *et al.* Changing policy and practice: making sense of national guidelines for osteoarthritis. *Soc Sci Med* 2014;**106**:101–9.

Porcheret M, Main C, Croft P, McKinley R, Hassell A, Dziedzic K. Development of a behaviour change intervention: a case study on the practical application of theory. *Implement Sci* 2014;**9**:42.

Uthman OA, van der Windt DA, Jordan JL, Dziedzic KS, Healey EL, Peat GM, Foster NE. Exercise for lower limb osteoarthritis: systematic review incorporating trial sequential analysis and network meta-analysis. *Br J Sports Med* 2014;**48**:1579.

Edwards JJ, Jordan KP, Peat G, Bedson J, Croft PR, Hay EM, Dziedzic KS. Quality of care for OA: the effect of a point-of-care consultation recording template. *Rheumatology* 2015;**54**:844–53.

Edwards JJ, Khanna M, Jordan KP, Jordan JL, Bedson J, Dziedzic KS. Quality indicators for the primary care of osteoarthritis: a systematic review. *Ann Rheum Dis* 2015;**74**:490–8.

Keeley T, Al-Janabi H, Nicholls E, Foster NE, Jowett S, Coast J. A longitudinal assessment of the responsiveness of the ICECAP-A in a randomised controlled trial of a knee pain intervention. *Qual Life Res* 2015;**24**:2319–31.

Morden A, Brooks L, Jinks C, Porcheret M, Ong BN, Dziedzic K. Research 'push', long term-change, and general practice. *J Health Organ Manag* 2015;**29**:798–821.

Morden A, Ong BN, Brooks L, Jinks C, Porcheret M, Edwards JJ, Dziedzic KS. Introducing Evidence Through Research 'Push': Using Theory and Qualitative Methods. *Qual Health Res* 2015;**25**:1560–75.

Østerås N, Jordan KP, Clausen B, Cordeiro C, Dziedzic K, Edwards J, *et al.* Self-reported quality care for knee osteoarthritis: comparisons across Denmark, Norway, Portugal and the UK. *RMD Open* 2015;**1**:e000136.

Quicke JG, Foster NE, Thomas MJ, Holden MA. Is long-term physical activity safe for older adults with knee pain?: a systematic review. *Osteoarthr Cartil* 2015;**23**:1445–56.

Blackburn S, Higginbottom A, Taylor R, Bird J, Østerås N, Hagen KB, *et al.* Patient-reported quality indicators for osteoarthritis: a patient and public generated self-report measure for primary care. *Res Involv Engagem* 2016;**2**:5.

Dziedzic KS, French S, Davis AM, Geelhoed E, Porcheret M. Implementation of musculoskeletal Models of Care in primary care settings: Theory, practice, evaluation and outcomes for musculoskeletal health in high-income economies. *Best Pract Res Clin Rheumatol* 2016;**30**:375–97.

Finney A, Healey E, Jordan JL, Ryan S, Dziedzic KS. Multidisciplinary approaches to managing osteoarthritis in multiple joint sites: a systematic review. *BMC Musculoskelet Disord* 2016;**17**:266.

Healey EL, Main CJ, Ryan S, McHugh GA, Porcheret M, Finney AG, *et al.* A nurse-led clinic for patients consulting with osteoarthritis in general practice: development and impact of training in a cluster randomised controlled trial. *BMC Fam Pract* 2016;**17**:173.

Jordan KP, Tan V, Edwards JJ, Chen Y, Englund M, Hubertsson J, *et al.* Influences on the decision to use an osteoarthritis diagnosis in primary care: a cohort study with linked survey and electronic health record data. *Osteoarthr Cartil* 2016;**24**:786–93.

Barnett LA, Lewis M, Mallen CD, Peat G. Applying quantitative bias analysis to estimate the plausible effects of selection bias in a cluster randomised controlled trial: secondary analysis of the Primary care Osteoarthritis Screening Trial (POST). *Trials* 2017;**18**:585.

Finney A, Dziedzic KS, Lewis M, Healey E. Multisite peripheral joint pain: a cross-sectional study of prevalence and impact on general health, quality of life, pain intensity and consultation behaviour. *BMC Musculoskelet Disord* 2017;**18**:535.

Holden MA, Whittle R, Healey EL, Hill S, Mullis R, Roddy E, Sowden G, Tooth S, Foster NE. Content and evaluation of the benefits of effective exercise for older adults with knee pain trial physiotherapist training program. *Arch Phys Med Rehabil* 2017;**98**:866–73.

Jackson H, Barnett LA, Jordan KP, Dziedzic KS, Cottrell E, Finney AG, *et al.* Patterns of routine primary care for osteoarthritis in the UK: a cross-sectional electronic health records study. *BMJ Open* 2017;**7**:e019694.

Jordan KP, Edwards JJ, Porcheret M, Healey EL, Jinks C, Bedson J, *et al.* Effect of a model consultation informed by guidelines on recorded quality of care of osteoarthritis (MOSAICS): a cluster randomised controlled trial in primary care. *Osteoarthr Cartil* 2017;**25**:1588–97.

Mallen CD, Nicholl BI, Lewis M, Bartlam B, Green D, Jowett S, *et al.* The effects of implementing a point-of-care electronic template to prompt routine anxiety and depression screening in patients consulting for osteoarthritis (the Primary Care Osteoarthritis Trial): a cluster randomised trial in primary care. *PLOS Med* 2017;**14**:e1002273.

Quicke JG, Foster NE, Ogollah RO, Croft PR, Holden MA. Relationship Between Attitudes and Beliefs and Physical Activity in Older Adults With Knee Pain: Secondary Analysis of a Randomized Controlled Trial. *Arthritis Care Res* 2017;**69**:1192–200.

Dziedzic KS, Healey EL, Porcheret M, Afolabi EK, Lewis M, Morden A, *et al.* Implementing core nice guidelines for osteoarthritis in primary care with a model consultation (MOSAICS): a cluster randomised controlled trial. *Osteoarthr Cartil* 2018;**26**:43–53.

Kigozi J, Jowett S, Nicholl BI, Lewis M, Bartlam B, Green D, *et al.* Cost-utility analysis of routine anxiety and depression screening in patients consulting for osteoarthritis: results from the POST trial [published online ahead of print April 2 2018]. *Arthritis Care Res* 2018.

Kigozi J, Jowett S, Nicholls E, Tooth S, Hay EM, Foster NE, *et al.* Cost-utility analysis of interventions to improve effectiveness of exercise therapy for adults with knee osteoarthritis: the BEEP trial [published online ahead of print June 6 2018]. *Rheumatol Adv Pract* 2018.

Oppong R, Jowett S, Lewis M, Clarkson K, Paskins Z, Croft P, *et al.* Cost-effectiveness of a model consultation to support self-management in patients with osteoarthritis. *Rheumatology* 2018;**57**:1056–63.

Porcheret M, Main C, Croft P, Dziedzic K. Enhancing delivery of osteoarthritis care in the general practice consultation: evaluation of a behaviour change intervention. *BMC Fam Pract* 2018;**19**:26.

Quicke JG, Foster NE, Croft PR, Ogollah RO, Holden MA. Change in physical activity level and clinical outcomes in older adults with knee pain: a secondary analysis from a randomised controlled trial. *BMC Musculoskelet Disord* 2018;**19**:59.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Data sharing statement

The authors' data archiving position is stated in our Standard Operating Procedure on Archiving and Destruction (www.keele.ac.uk/pchs/intranet/sops/sop17/).[354]

References

1. Cross M, Smith E, Hoy D, Nolte S, Ackerman I, Fransen M, *et al.* The global burden of hip and knee osteoarthritis: estimates from the global burden of disease 2010 study. *Ann Rheum Dis* 2014;**73**:1323–30. https://doi.org/10.1136/annrheumdis-2013-204763
2. Arthritis Research UK. *State of Musculoskeletal Health 2017*. URL: www.arthritisresearchuk.org/arthritis-information/data-and-statistics/state-of-musculoskeletal-health.aspx (accessed 11 May 2017).
3. Arthritis Research UK Primary Care Centre, Keele University. *Musculoskeletal Matters (Bulletin 1): What do General Practitioners see?* 2009. URL: www.keele.ac.uk/media/keeleuniversity/ri/primarycare/bulletins/MusculoskeletalMatters1.pdf (accessed 11 May 2017).
4. Arthritis Research UK Primary Care Centre, Keele University. *Musculoskeletal Matters (Bulletin 2): Consultations for Selected Diagnoses and Regional Problems*; 2010. URL: www.keele.ac.uk/media/keeleuniversity/ri/primarycare/bulletins/MusculoskeletalMatters2.pdf (accessed 11 May 2017).
5. Jagger C, Matthews R, Spiers N, Brayne C, Comas-Herrera A, Robinson T, *et al. Compression or expansion of disability?* 2006. URL: www.kingsfund.org.uk/sites/files/kf/compression-expansion-disability-wanless-research-paper-carol-jagger2006.pdf (accessed 11 May 2017).
6. Mann C, Gooberman-Hill R. Health care provision for osteoarthritis: concordance between what patients would like and what health professionals think they should have. *Arthritis Care Res* 2011;**63**:963–72. https://doi.org/10.1002/acr.20459
7. Woolhead GM, Donovan JL, Dieppe PA. Outcomes of total knee replacement: a qualitative study. *Rheumatology* 2005;**44**:1032–7. https://doi.org/10.1093/rheumatology/keh674
8. Anon. Recommendations for the medical management of osteoarthritis of the hip and knee: 2000 update. American College of Rheumatology Subcommittee on Osteoarthritis Guidelines. *Arthritis Rheum* 2000;**43**:1905–15. https://doi.org/10.1002/1529-0131(200009)43:9<1905::AID-ANR1>3.0.CO;2-P
9. Jordan KM, Arden NK, Doherty M, Bannwarth B, Bijlsma JW, Dieppe P, *et al.* EULAR Recommendations 2003: an evidence based approach to the management of knee osteoarthritis: Report of a Task Force of the Standing Committee for International Clinical Studies Including Therapeutic Trials (ESCISIT). *Ann Rheum Dis* 2003;**62**:1145–55. https://doi.org/10.1136/ard.2003.011742
10. Zhang W, Doherty M, Leeb BF, Alekseeva L, Arden NK, Bijlsma JW, *et al.* EULAR evidence based recommendations for the management of hand osteoarthritis: report of a Task Force of the EULAR Standing Committee for International Clinical Studies Including Therapeutics (ESCISIT). *Ann Rheum Dis* 2007;**66**:377–88. https://doi.org/10.1136/ard.2006.062091
11. Zhang W, Doherty M, Leeb BF, Alekseeva L, Arden NK, Bijlsma JW, *et al.* EULAR evidence-based recommendations for the diagnosis of hand osteoarthritis: report of a task force of ESCISIT. *Ann Rheum Dis* 2009;**68**:8–17. https://doi.org/10.1136/ard.2007.084772
12. Zhang W, Moskowitz RW, Nuki G, Abramson S, Altman RD, Arden N, *et al.* OARSI recommendations for the management of hip and knee osteoarthritis, part I: critical appraisal of existing treatment guidelines and systematic review of current research evidence. *Osteoarthr Cartil* 2007;**15**:981–1000. https://doi.org/10.1016/j.joca.2007.06.014
13. Zhang W, Moskowitz RW, Nuki G, Abramson S, Altman RD, Arden N, *et al.* OARSI recommendations for the management of hip and knee osteoarthritis, Part II: OARSI evidence-based, expert consensus guidelines. *Osteoarthr Cartil* 2008;**16**:137–62. https://doi.org/10.1016/j.joca.2007.12.013

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

14. Zhang W, Doherty M, Peat G, Bierma-Zeinstra MA, Arden NK, Bresnihan B, *et al.* EULAR evidence-based recommendations for the diagnosis of knee osteoarthritis. *Ann Rheum Dis* 2010;**69**:483–9. https://doi.org/10.1136/ard.2009.113100

15. National Institute for Health and Clinical Excellence. *Osteoarthritis: The Care And Management Of Osteoarthritis In Adults*. Clinical Guideline CG59. London: NICE; 2008.

16. van Dijk GM, Veenhof C, Schellevis F, Hulsmans H, Bakker JP, Arwert H, *et al.* Comorbidity, limitations in activities and pain in patients with osteoarthritis of the hip or knee. *BMC Musculoskelet Disord* 2008;**9**:95. https://doi.org/10.1186/1471-2474-9-95

17. Kadam UT, Jordan K, Croft PR. Clinical comorbidity in patients with osteoarthritis: a case-control study of general practice consulters in England and Wales. *Ann Rheum Dis* 2004;**63**:408–14. https://doi.org/10.1136/ard.2003.007526

18. Lin EH, Katon W, Von Korff M, Tang L, Williams JW, Kroenke K, *et al.* Effect of improving depression care on pain and functional outcomes among older adults with arthritis: a randomized controlled trial. *JAMA* 2003;**290**:2428–9. https://doi.org/10.1001/jama.290.18.2428

19. Kennedy A, Rogers A, Bower P. Support for self care for patients with chronic disease. *BMJ* 2007;**335**:968–70. https://doi.org/10.1136/bmj.39372.540903.94

20. Kennedy AP, Nelson E, Reeves D, Richardson G, Roberts C, Robinson A, *et al.* A randomised controlled trial to assess the effectiveness and cost of a patient orientated self management approach to chronic inflammatory bowel disease. *Gut* 2004;**53**:1639–45. https://doi.org/10.1136/gut.2003.034256

21. Robinson A, Lee V, Kennedy A, Middleton L, Rogers A, Thompson DG, Reeves D. A randomised controlled trial of self-help interventions in patients with a primary care diagnosis of irritable bowel syndrome. *Gut* 2006;**55**:643–8. https://doi.org/10.1136/gut.2004.062901

22. Jinks C, Ong BN, Richardson J. A mixed methods study to investigate needs assessment for knee pain and disability: population and individual perspectives. *BMC Musculoskelet Disord* 2007;**8**:59. https://doi.org/10.1186/1471-2474-8-59

23. Steel N, Bachmann M, Maisey S, Shekelle P, Breeze E, Marmot M, Melzer D. Self reported receipt of care consistent with 32 quality indicators: national population survey of adults aged 50 or more in England. *BMJ* 2008;**337**:a957. https://doi.org/10.1136/bmj.a957

24. Porcheret M, Jordan K, Jinks C, Croft P, Primary Care Rheumatology Society. Primary care treatment of knee pain – a survey in older adults. *Rheumatology* 2007;**46**:1694–700. https://doi.org/10.1093/rheumatology/kem232

25. Department of Health. *The NHS Improvement Plan: Putting People at the Heart of Public Services*. London: DH; 2004.

26. Kralik D, Koch T, Price K, Howard N. Chronic illness self-management: taking action to create order. *J Clin Nurs* 2004;**13**:259–67. https://doi.org/10.1046/j.1365-2702.2003.00826.x

27. Thorne S, Paterson B, Russell C. The structure of everyday self-care decision making in chronic illness. *Qual Health Res* 2003;**13**:1337–52. https://doi.org/10.1177/1049732303258039

28. Townsend A, Wyke S, Hunt K. Self-managing and managing self: practical and moral dilemmas in accounts of living with chronic illness. *Chronic Illn* 2006;**2**:185–94. https://doi.org/10.1177/17423953060020031301

29. Robison JI, Rogers MA. Adherence to exercise programmes. Recommendations. *Sports Med* 1994;**17**:39–52. https://doi.org/10.2165/00007256-199417010-00004

30. Sullivan T, Allegrante JP, Peterson MG, Kovar PA, MacKenzie CR. One-year follow-up of patients with osteoarthritis of the knee who participated in a program of supervised fitness walking and supportive patient education. *Arthritis Care Res* 1998;**11**:228–33. https://doi.org/10.1002/art.1790110403

31. Williams NH, Hendry M, France B, Lewis R, Wilkinson C. Effectiveness of exercise-referral schemes to promote physical activity in adults: systematic review. *Br J Gen Pract* 2007;**57**:979–86. https://doi.org/10.3399/096016407782604866

32. Roddy E, Zhang W, Doherty M, Arden NK, Barlow J, Birrell F, *et al.* Evidence-based recommendations for the role of exercise in the management of osteoarthritis of the hip or knee – the MOVE consensus. *Rheumatology* 2005;**44**:67–73. https://doi.org/10.1093/rheumatology/keh399

33. Fransen M, McConnell S. Exercise for osteoarthritis of the knee. *Cochrane Database Syst Rev* 2008;**4**:CD004376. https://doi.org/10.1002/14651858.CD004376.pub2

34. Hay EM, Foster NE, Thomas E, Peat G, Phelan M, Yates HE, *et al.* Effectiveness of community physiotherapy and enhanced pharmacy review for knee pain in people aged over 55 presenting to primary care: pragmatic randomised trial. *BMJ* 2006;**333**:995. https://doi.org/10.1136/bmj.38977.590752.0B

35. Gureje O, Von Korff M, Simon GE, Gater R. Persistent pain and well-being: a World Health Organization Study in Primary Care. *JAMA* 1998;**280**:147–51. https://doi.org/10.1001/jama.280.2.147

36. Bair MJ, Robinson RL, Katon W, Kroenke K. Depression and pain comorbidity: a literature review. *Arch Intern Med* 2003;**163**:2433–45. https://doi.org/10.1001/archinte.163.20.2433

37. Arnow BA, Hunkeler EM, Blasey CM, Lee J, Constantino MJ, Fireman B, *et al.* Comorbid depression, chronic pain, and disability in primary care. *Psychosom Med* 2006;**68**:262–8. https://doi.org/10.1097/01.psy.0000204851.15499.fc

38. Rosemann T, Backenstrass M, Joest K, Rosemann A, Szecsenyi J, Laux G. Predictors of depression in a sample of 1,021 primary care patients with osteoarthritis. *Arthritis Rheum* 2007;**57**:415–22. https://doi.org/10.1002/art.22624

39. Axford J, Heron C, Ross F, Victor CR. Management of knee osteoarthritis in primary care: pain and depression are the major obstacles. *J Psychosom Res* 2008;**64**:461–7. https://doi.org/10.1016/j.jpsychores.2007.11.009

40. Rose G. *The Strategy of Preventive Medicine*. Oxford: Oxford University Press; 1982.

41. McHugh GA, Silman AJ, Luker KA. Quality of care for people with osteoarthritis: a qualitative study. *J Clin Nurs* 2007;**16**:168–76. https://doi.org/10.1111/j.1365-2702.2007.01885.x

42. Gignac MA, Davis AM, Hawker G, Wright JG, Mahomed N, Fortin PR, Badley EM. "What do you expect? You're just getting older": A comparison of perceived osteoarthritis-related and aging-related health experiences in middle- and older-age adults. *Arthritis Rheum* 2006;**55**:905–12. https://doi.org/10.1002/art.22338

43. Ong BN, Jinks C, Morden A. The hard work of self-management: living with chronic knee pain. *Int J Qual Stud Health Well-being* 2011;**6**:7035. https://doi.org/10.3402/qhw.v6i3.7035

44. Hill S, Dziedzic K, Thomas E, Baker SR, Croft P. The illness perceptions associated with health and behavioural outcomes in people with musculoskeletal hand problems: findings from the North Staffordshire Osteoarthritis Project (NorStOP). *Rheumatology* 2007;**46**:944–51. https://doi.org/10.1093/rheumatology/kem015

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

45. Takahashi H, Nakajima M, Ozaki K, Tanaka T, Kamatani N, Ikegawa S. Prediction model for knee osteoarthritis based on genetic and clinical information. *Arthritis Res Ther* 2010;**12**:R187. https://doi.org/10.1186/ar3157

46. Kerkhof HJ, Bierma-Zeinstra SM, Arden NK, Metrustry S, Castano-Betancourt M, Hart DJ, *et al.* Prediction model for knee osteoarthritis incidence, including clinical, genetic and biochemical risk factors. *Ann Rheum Dis* 2014;**73**:2116–21. https://doi.org/10.1136/annrheumdis-2013-203620

47. French HP, Galvin R, Cusack T, McCarthy GM. Predictors of short-term outcome to exercise and manual therapy for people with hip osteoarthritis. *Phys Ther* 2014;**94**:31–9. https://doi.org/10.2522/ptj.20130173

48. Wright AA, Cook CE, Flynn TW, Baxter GD, Abbott JH. Predictors of response to physical therapy intervention in patients with primary hip osteoarthritis. *Phys Ther* 2011;**91**:510–24. https://doi.org/10.2522/ptj.20100171

49. Jinks C, Jordan KP, Blagojevic M, Croft P. Predictors of onset and progression of knee pain in adults living in the community. A prospective study. *Rheumatology* 2008;**47**:368–74. https://doi.org/10.1093/rheumatology/kem374

50. Thomas E, Peat G, Mallen C, Wood L, Lacey R, Duncan R, Croft P. Predicting the course of functional limitation among older adults with knee pain: do local signs, symptoms and radiographs add anything to general indicators? *Ann Rheum Dis* 2008;**67**:1390–8. https://doi.org/10.1136/ard.2007.080945

51. Zhang W, McWilliams DF, Ingham SL, Doherty SA, Muthuri S, Muir KR, Doherty M. Nottingham knee osteoarthritis risk prediction models. *Ann Rheum Dis* 2011;**70**:1599–604. https://doi.org/10.1136/ard.2011.149807

52. Yusuf E, Bijsterbosch J, Slagboom PE, Kroon HM, Rosendaal FR, Huizinga TW, Kloppenburg M. Association between several clinical and radiological determinants with long-term clinical progression and good prognosis of lower limb osteoarthritis. *PLOS ONE* 2011;**6**:e25426. https://doi.org/10.1371/journal.pone.0025426

53. Birrell FN. Patterns of joint pain: lessons from epidemiology. *Rheumatology* 2004;**43**:408–9. https://doi.org/10.1093/rheumatology/keh129

54. Peat G, Thomas E, Wilkie R, Croft P. Multiple joint pain and lower extremity disability in middle and old age. *Disabil Rehabil* 2006;**28**:1543–9. https://doi.org/10.1080/09638280600646250

55. Keenan AM, Tennant A, Fear J, Emery P, Conaghan PG. Impact of multiple joint problems on daily living tasks in people in the community over age fifty-five. *Arthritis Rheum* 2006;**55**:757–64. https://doi.org/10.1002/art.22239

56. Segal L, Day SE, Chapman AB, Osborne RH. Can we reduce disease burden from osteoarthritis? *Med J Aust* 2004;**180**(Suppl. 5):11–17.

57. Andrews G, Simonella L, Lapsley H, Sanderson K, March L. Evidence-based medicine is affordable: the cost-effectiveness of current compared with optimal treatment in rheumatoid and osteoarthritis. *J Rheumatol* 2006;**33**:671–80.

58. Gupta S, Hawker GA, Laporte A, Croxford R, Coyte PC. The economic burden of disabling hip and knee osteoarthritis (OA) from the perspective of individuals living with this condition. *Rheumatology* 2005;**44**:1531–7. https://doi.org/10.1093/rheumatology/kei049

59. Wulff J. *Modelling Clinical Outcomes and Cost-effectiveness of Primary Care Interventions for Osteoarthritis Using Prediction and Decision Models*. PhD thesis. Keele: Keele University; 2013.

60. Thomas E, Wilkie R, Peat G, Hill S, Dziedzic K, Croft P. The North Staffordshire Osteoarthritis Project–NorStOP: prospective, 3-year study of the epidemiology and management of clinical osteoarthritis in a general population of older adults. *BMC Musculoskelet Disord* 2004;**13**;5:2.

61. Bellamy N, Kirwan J, Boers M, Brooks P, Strand V, Tugwell P, *et al.* Recommendations for a core set of outcome measures for future phase III clinical trials in knee, hip, and hand osteoarthritis. Consensus development at OMERACT III. *J Rheumatol* 1997;**24**:799–802.

62. Bellamy N. *WOMAC Osteoarthritis Index. A User's Guide*. London, ON: University of Western Ontario; 1996.

63. Bellamy N, Campbell J, Haraoui B, Buchbinder R, Hobby K, Roth JH, MacDermid JC. Dimensionality and clinical importance of pain and disability in hand osteoarthritis: Development of the Australian/Canadian (AUSCAN) Osteoarthritis Hand Index. *Osteoarthr Cartil* 2002;**10**:855–62. https://doi.org/10.1053/joca.2002.0837

64. Garrow AP, Papageorgiou AC, Silman AJ, Thomas E, Jayson MI, Macfarlane GJ. Development and validation of a questionnaire to assess disabling foot pain. *Pain* 2000;**85**:107–13. https://doi.org/10.1016/S0304-3959(99)00263-8

65. Menz HB, Tiedemann A, Kwan MM, Plumb K, Lord SR. Foot pain in community-dwelling older people: an evaluation of the Manchester Foot Pain and Disability Index. *Rheumatology* 2006;**45**:863–7. https://doi.org/10.1093/rheumatology/kel002

66. Zelman DC, Hoffman DL, Seifeldin R, Dukes EM. Development of a metric for a day of manageable pain control: derivation of pain severity cut-points for low back pain and osteoarthritis. *Pain* 2003;**106**:35–42. https://doi.org/10.1016/S0304-3959(03)00274-4

67. Ware JE Jr., Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Medical Care* 1992;**30**:473–83. https://doi.org/10.1097/00005650-199206000-00002

68. Tu JV, Austin PC, Walld R, Roos L, Agras J, McDonald KM. Development and validation of the Ontario acute myocardial infarction mortality prediction rules. *J Am Coll Cardiol* 2001;**37**:992–7. https://doi.org/10.1016/S0735-1097(01)01109-3

69. Zou G. A modified poisson regression approach to prospective studies with binary data. *Am J Epidemiol* 2004;**159**:702–6. https://doi.org/10.1093/aje/kwh090

70. Royston P, Altman DG, Sauerbrei W. Dichotomizing continuous predictors in multiple regression: a bad idea. *Stat Med* 2006;**25**:127–41. https://doi.org/10.1002/sim.2331

71. Smit F, Ederveen A, Cuijpers P, Deeg D, Beekman A. Opportunities for cost-effective prevention of late-life depression: an epidemiological approach. *Arch Gen Psychiatry* 2006;**63**:290–6. https://doi.org/10.1001/archpsyc.63.3.290

72. Greenland S, Drescher K. Maximum likelihood estimation of the attributable fraction from logistic models. *Biometrics* 1993;**49**:865–72. https://doi.org/10.2307/2532206

73. Muller S. *The Measurement of Locomotor Disability in Epidemiological Studies*. PhD thesis. Keele: Keele University; 2010.

74. Rockhill B, Newman B, Weinberg C. Use and misuse of population attributable fractions. *Am J Public Health* 1998;**88**:15–19. https://doi.org/10.2105/AJPH.88.1.15

75. Muller S, Thomas E, Peat G. The effect of changes in lower limb pain on the rate of progression of locomotor disability in middle and old age: evidence from the NorStOP cohort with 6-year follow-up. *Pain* 2012;**153**:952–9. https://doi.org/10.1016/j.pain.2011.12.006

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

76. Shah RC, Buchman AS, Boyle PA, Leurgans SE, Wilson RS, Andersson GB, Bennett DA. Musculoskeletal pain is associated with incident mobility disability in community-dwelling elders. *J Gerontol A Biol Sci Med Sci* 2011;**66**:82–8. https://doi.org/10.1093/gerona/glq187

77. Hartvigsen J, Natvig B, Ferreira M. Is it all about a pain in the back? *Best Pract Res Clin Rheumatol* 2013;**27**:613–23. https://doi.org/10.1016/j.berh.2013.09.008

78. Steyerberg EW, Eijkemans MJ, Harrell FE, Habbema JD. Prognostic modelling with logistic regression analysis: a comparison of selection and estimation methods in small data sets. *Stat Med* 2000;**19**:1059–79. https://doi.org/10.1002/(SICI)1097-0258(20000430)19:8<1059::AID-SIM412>3.0.CO;2-0

79. Lacey RJ, Jordan KP, Croft PR. Does attrition during follow-up of a population cohort study inevitably lead to biased estimates of health status? *PLOS ONE* 2013;**8**:e83948. https://doi.org/10.1371/journal.pone.0083948

80. Hedges LV, Olkin I. *Statistical Methods for Meta-analysis*. Orlando, FL: Academic Press; 1985.

81. Cohen J. *Statistical Power Analysis for the Behavioural Sciences*. New York, NY: Academic Press; 1977.

82. Higgins JP, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, *et al.* The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 2011;**343**:d5928. https://doi.org/10.1136/bmj.d5928

83. Riley RD, Higgins JP, Deeks JJ. Interpretation of random effects meta-analyses. *BMJ* 2011;**342**:d549. https://doi.org/10.1136/bmj.d549

84. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997;**315**:629–34. https://doi.org/10.1136/bmj.315.7109.629

85. Bautch JC, Malone DG, Vailas AC. Effects of exercise on knee joints with osteoarthritis: a pilot study of biologic markers. *Arthritis Care Res* 1997;**10**:48–55. https://doi.org/10.1002/art.1790100108

86. Ettinger WH, Burns R, Messier SP, Applegate W, Rejeski WJ, Morgan T, *et al.* A randomized trial comparing aerobic exercise and resistance exercise with a health education program in older adults with knee osteoarthritis. The Fitness Arthritis and Seniors Trial (FAST). *JAMA* 1997;**277**:25–31 https://doi.org/10.1001/jama.1997.03540250033028.

87. van Baar ME, Dekker J, Oostendorp RA, Bijl D, Voorn TB, Lemmens JA, Bijlsma JW. The effectiveness of exercise therapy in patients with osteoarthritis of the hip or knee: a randomized clinical trial. *J Rheumatol* 1998;**25**:2432–9.

88. Maurer BT, Stern AG, Kinossian B, Cook KD, Schumacher HR. Osteoarthritis of the knee: isokinetic quadriceps exercise versus an educational intervention. *Arch Phys Med Rehabil* 1999;**80**:1293–9. https://doi.org/10.1016/S0003-9993(99)90032-1

89. O'Reilly SC, Muir KR, Doherty M. Effectiveness of home exercise on pain and disability from osteoarthritis of the knee: a randomised controlled trial. *Ann Rheum Dis* 1999;**58**:15–19. https://doi.org/10.1136/ard.58.1.15

90. Péloquin L, Bravo G, Gauthier P, Lacombe G, Billiard JS. Effects of a cross-training exercise program in persons with osteoarthritis of the knee a randomized controlled trial. *J Clin Rheumatol* 1999;**5**:126–36. https://doi.org/10.1097/00124743-199906000-00004

91. Deyle GD, Henderson NE, Matekel RL, Ryder MG, Garber MB, Allison SC. Effectiveness of manual physical therapy and exercise in osteoarthritis of the knee. A randomized, controlled trial. *Ann Intern Med* 2000;**132**:173–81. https://doi.org/10.7326/0003-4819-132-3-200002010-00002

92. Hopman-Rock M, Westhoff MH. The effects of a health educational and exercise program for older adults with osteoarthritis for the hip or knee. *J Rheumatol* 2000;**27**:1947–54.

93. Baker KR, Nelson ME, Felson DT, Layne JE, Sarno R, Roubenoff R. The efficacy of home based progressive strength training in older adults with knee osteoarthritis: a randomized controlled trial. *J Rheumatol* 2001;**28**:1655–65.

94. Fransen M, Crosbie J, Edmonds J. Physical therapy is effective for patients with osteoarthritis of the knee: a randomized controlled clinical trial. *J Rheumatol* 2001;**28**:156–64.

95. Halbert J, Crotty M, Weller D, Ahern M, Silagy C. Primary care-based physical activity programs: effectiveness in sedentary older patients with osteoarthritis symptoms. *Arthritis Rheum* 2001;**45**:228–34. https://doi.org/10.1002/1529-0131(200106)45:3<228::AID-ART253>3.0.CO;2-2

96. Patrick DL, Ramsey SD, Spencer AC, Kinne S, Belza B, Topolski TD. Economic evaluation of aquatic exercise for persons with osteoarthritis. *Med Care* 2001;**39**:413–24. https://doi.org/10.1097/00005650-200105000-00002

97. Belza B, Topolski T, Kinne S, Patrick DL, Ramsey SD. Does adherence make a difference? Results from a community-based aquatic exercise program. *Nurs Res* 2002;**51**:285–91. https://doi.org/10.1097/00006199-200209000-00003

98. Thomas KS, Muir KR, Doherty M, Jones AC, O'Reilly SC, Bassey EJ. Home based exercise programme for knee pain and knee osteoarthritis: randomised controlled trial. *BMJ* 2002;**325**:752. https://doi.org/10.1136/bmj.325.7367.752

99. Topp R, Woolley S, Hornyak J, Khuder S, Kahaleh B. The effect of dynamic versus isometric resistance training on pain and functioning among adults with osteoarthritis of the knee. *Arch Phys Med Rehabil* 2002;**83**:1187–95. https://doi.org/10.1053/apmr.2002.33988

100. Foley A, Halbert J, Hewitt T, Crotty M. Does hydrotherapy improve strength and physical function in patients with osteoarthritis – a randomised controlled trial comparing a gym based and a hydrotherapy based strengthening programme. *Ann Rheum Dis* 2003;**62**:1162–7. https://doi.org/10.1136/ard.2002.005272

101. Quilty B, Tucker M, Campbell R, Dieppe P. Physiotherapy, including quadriceps exercises and patellar taping, for knee osteoarthritis with predominant patello-femoral joint involvement: randomized controlled trial. *J Rheumatol* 2003;**30**:1311–17.

102. Talbot LA, Gaines JM, Huynh TN, Metter EJ. A home-based pedometer-driven walking program to increase physical activity in older adults with osteoarthritis of the knee: a preliminary study. *J Am Geriatr Soc* 2003;**51**:387–92. https://doi.org/10.1046/j.1532-5415.2003.51113.x

103. Hughes SL, Seymour RB, Campbell R, Pollak N, Huber G, Sharma L. Impact of the fit and strong intervention on older adults with osteoarthritis. *Gerontologist* 2004;**44**:217–28. https://doi.org/10.1093/geront/44.2.217

104. Keefe FJ, Blumenthal J, Baucom D, Affleck G, Waugh R, Caldwell DS, *et al.* Effects of spouse-assisted coping skills training and exercise training in patients with osteoarthritic knee pain: a randomized controlled study. *Pain* 2004;**110**:539–49. https://doi.org/10.1016/j.pain.2004.03.022

105. Lin SY, Davey RC, Cochrane T. Community rehabilitation for older adults with osteoarthritis of the lower limb: a controlled clinical trial. *Clin Rehabil* 2004;**18**:92–101. https://doi.org/10.1191/0269215504cr706oa

106. Messier SP, Loeser RF, Miller GD, Morgan TM, Rejeski WJ, Sevick MA, *et al.* Exercise and dietary weight loss in overweight and obese older adults with knee osteoarthritis: the Arthritis, Diet, and Activity Promotion Trial. *Arthritis Rheum* 2004;**50**:1501–10. https://doi.org/10.1002/art.20256

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

107. Ravaud P, Giraudeau B, Logeart I, Larguier JS, Rolland D, Treves R, *et al.* Management of osteoarthritis (OA) with an unsupervised home based exercise programme and/or patient administered assessment tools. A cluster randomised controlled trial with a 2x2 factorial design. *Ann Rheum Dis* 2004;**63**:703–8. https://doi.org/10.1136/ard.2003.009803

108. Bennell KL, Hinman RS, Metcalf BR, Buchbinder R, McConnell J, McColl G, *et al.* Efficacy of physiotherapy management of knee joint osteoarthritis: a randomised, double blind, placebo controlled trial. *Ann Rheum Dis* 2005;**64**:906–12. https://doi.org/10.1136/ard.2004.026526

109. Cochrane T, Davey RC, Matthes Edwards SM. Randomised controlled trial of the cost-effectiveness of water-based therapy for lower limb osteoarthritis. *Health Technol Assess* 2005;**9**(31). https://doi.org/10.3310/hta9310

110. Tak E, Staats P, Van Hespen A, Hopman-Rock M. The effects of an exercise program for older adults with osteoarthritis of the hip. *J Rheumatol* 2005;**32**:1106–13.

111. Mikesky AE, Mazzuca SA, Brandt KD, Perkins SM, Damush T, Lane KA. Effects of strength training on the incidence and progression of knee osteoarthritis. *Arthritis Rheum* 2006;**55**:690–9. https://doi.org/10.1002/art.22245

112. Fransen M, Nairn L, Winstanley J, Lam P, Edmonds J. The Physical Activity for Osteoarthritis Management (PAFORM) study. A randomised controlled clinical trial evaluating hydrotherapy and Tai Chi classes. *Arthritis Care Res* 2007;**57**:407–14. https://doi.org/10.1002/art.22621

113. Hinman RS, Heywood SE, Day AR. Aquatic physical therapy for hip and knee osteoarthritis: results of a single-blind randomized controlled trial. *Phys Ther* 2007;**87**:32–43. https://doi.org/10.2522/ptj.20060006

114. Wang TJ, Belza B, Elaine Thompson F, Whitney JD, Bennett K. Effects of aquatic exercise on flexibility, strength and aerobic fitness in adults with osteoarthritis of the hip or knee. *J Adv Nurs* 2007;**57**:141–52. https://doi.org/10.1111/j.1365-2648.2006.04102.x

115. Superio-Cabuslay E, Ward MM, Lorig KR. Patient education interventions in osteoarthritis and rheumatoid arthritis: a meta-analytic comparison with nonsteroidal antiinflammatory drug treatment. *Arthritis Care Res* 1996;**9**:292–301. https://doi.org/10.1002/1529-0131(199608)9:4<292::AID-ANR1790090414>3.0.CO;2-4

116. Zhang W, Jones A, Doherty M. Does paracetamol (acetaminophen) reduce the pain of osteoarthritis? A meta-analysis of randomised controlled trials. *Ann Rheum Dis* 2004;**63**:901–7. https://doi.org/10.1136/ard.2003.018531

117. Towheed TE. Pennsaid therapy for osteoarthritis of the knee: a systematic review and metaanalysis of randomized controlled trials. *J Rheumatol* 2006;**33**:567–73.

118. Eccles M, Freemantle N, Mason J. North of England evidence based guideline development project: summary guideline for non-steroidal anti-inflammatory drugs versus basic analgesia in treating the pain of degenerative arthritis. The North of England Non-Steroidal Anti-Inflammatory Drug Guideline Development Group. *BMJ* 1998;**317**:526–30. https://doi.org/10.1136/bmj.317.7157.526

119. Biswal S, Medhi B, Pandhi P. Longterm efficacy of topical nonsteroidal antiinflammatory drugs in knee osteoarthritis: metaanalysis of randomized placebo controlled clinical trials. *J Rheumatol* 2006;**33**:1841–4.

120. Fransen M, McConnell S, Harmer AR, Van der Esch M, Simic M, Bennell KL. Exercise for osteoarthritis of the knee. *Cochrane Database Syst Rev* 2015;**1**:CD004376. https://doi.org/10.1002/14651858.CD004376.pub3

121. Jansen MJ, Viechtbauer W, Lenssen AF, Hendriks EJ, de Bie RA. Strength training alone, exercise therapy alone, and exercise therapy with passive manual mobilisation each reduce pain and disability in people with knee osteoarthritis: a systematic review. *J Physiother* 2011;**57**:11–20. https://doi.org/10.1016/S1836-9553(11)70002-9

122. Uthman OA, van der Windt DA, Jordan JL, Dziedzic KS, Healey EL, Peat GM, *et al.* Exercise for lower limb osteoarthritis: systematic review incorporating trial sequential analysis and network meta-analysis. *BMJ* 2013;**347**:f5555. https://doi.org/10.1136/bmj.f5555

123. Hurley MV, Walsh NE, Mitchell HL, Pimm TJ, Patel A, Williamson E, *et al.* Clinical effectiveness of a rehabilitation program integrating exercise, self-management, and active coping strategies for chronic knee pain: a cluster randomized trial. *Arthritis Rheum* 2007;**57**:1211–19. https://doi.org/10.1002/art.22995

124. Hurley MV, Walsh NE, Mitchell HL, Pimm TJ, Williamson E, Jones RH, *et al.* Economic evaluation of a rehabilitation program integrating exercise, self-management, and active coping strategies for chronic knee pain. *Arthritis Rheum* 2007;**57**:1220–9. https://doi.org/10.1002/art.23011

125. Loveman E, Copley VR, Colquitt J, Scott DA, Clegg A, Jones J, *et al.* The clinical effectiveness and cost-effectiveness of treatments for idiopathic pulmonary fibrosis: a systematic review and economic evaluation. *Health Technol Assess* 2015;**19**(20). https://doi.org/10.3310/hta19200

126. Joint Formulary Committee. *British National Formulary*. 64 edn. London: BMJ and Pharmaceutical Press; 2010.

127. Whitehurst DG, Bryan S, Hay EM, Thomas E, Young J, Foster NE. Cost-effectiveness of acupuncture care as an adjunct to exercise-based physical therapy for osteoarthritis of the knee. *Phys Ther* 2011;**91**:630–41. https://doi.org/10.2522/ptj.20100239

128. Brazier JE, Roberts J. The estimation of a preference-based measure of health from the SF-12. *Med Care* 2004;**42**:851–9. https://doi.org/10.1097/01.mlr.0000135827.18610.0d

129. Her Majesty's Treasury. *The Green Book: Appraisal and Evaluation in Central Government*. London: Her Majesty's Treasury; 2003.

130. Barton P, Bryan S, Robinson S. Modelling in the economic evaluation of health care: selecting the appropriate approach. *J Health Serv Res Policy* 2004;**9**:110–18. https://doi.org/10.1258/135581904322987535

131. Dziedzic KS, Healey EL, Porcheret M, Ong BN, Main CJ, Jordan KP, *et al.* Implementing the NICE osteoarthritis guidelines: a mixed methods study and cluster randomised trial of a model osteoarthritis consultation in primary care – the Management of OsteoArthritis In ConsultationS (MOSAICS) study protocol. *Implement Sci* 2014;**9**:95. https://doi.org/10.1186/s13012-014-0095-y

132. Edwards JJ, Jordan KP, Peat G, Bedson J, Croft PR, Hay EM, *et al.* Quality of care for OA: the effect of a point-of-care consultation recording template. *Rheumatology* 2015;**54**:844–53. https://doi.org/10.1093/rheumatology/keu411

133. Jinks C, Jordan K, Ong BN, Croft P. A brief screening tool for knee pain in primary care (KNEST). 2. Results from a survey in the general population aged 50 and over. *Rheumatology* 2004;**43**:55–61. https://doi.org/10.1093/rheumatology/keg438

134. Porcheret M, Main C, Croft P, McKinley R, Hassell A, Dziedzic K. Development of a behaviour change intervention: a case study on the practical application of theory. *Implement Sci* 2014;**9**:42. https://doi.org/10.1186/1748-5908-9-42

135. Grime JC, Ong BN. Constructing osteoarthritis through discourse – a qualitative analysis of six patient information leaflets on osteoarthritis. *BMC Musculoskelet Disord* 2007;**8**:34. https://doi.org/10.1186/1471-2474-8-34

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

136. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M, Medical Research Council Guidance. Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ* 2008;**337**:a1655. https://doi.org/10.1136/bmj.a1655

137. May C, Finch T. Implementing, embedding, and integrating practices: an outline of normalization process theory. *Sociology* 2009;**43**:535–54. https://doi.org/10.1177/0038038509103208

138. May C, Finch T, Mair F, Ballini L, Dowrick C, Eccles M, *et al.* Understanding the implementation of complex interventions in health care: the normalization process model. *BMC Health Serv Res* 2007;**7**:148. https://doi.org/10.1186/1472-6963-7-148

139. May C. Towards a general theory of implementation. *Implement Sci* 2013;**8**:18. https://doi.org/10.1186/1748-5908-8-18

140. Research Institute for Primary Care and Health Sciences. *The Osteoarthritis (OA) e-template*. URL: www.keele.ac.uk/mrr/ (accessed 10 April 2017).

141. Jordan KP, Jöud A, Bergknut C, Croft P, Edwards JJ, Peat G, *et al.* International comparisons of the consultation prevalence of musculoskeletal conditions using population-based healthcare data from England and Sweden. *Ann Rheum Dis* 2014;**73**:212–18. https://doi.org/10.1136/annrheumdis-2012-202634

142. Grime J, Dudley B. Developing written information on osteoarthritis for patients: facilitating user involvement by exposure to qualitative research. *Health Expect* 2014;**17**:164–73. https://doi.org/10.1111/j.1369-7625.2011.00741.x

143. Porcheret M, Grime J, Main C, Dziedzic K. Developing a model osteoarthritis consultation: a Delphi consensus exercise. *BMC Musculoskelet Disord* 2013;**14**:25. https://doi.org/10.1186/1471-2474-14-25

144. Finney A, Porcheret M, Grime J, Jordan KP, Handy J, Healey E, *et al.* Defining the content of an opportunistic osteoarthritis consultation with primary health care professionals: a Delphi consensus study. *Arthritis Care Res* 2013;**65**:962–8. https://doi.org/10.1002/acr.21917

145. Edwards JJ, Khanna M, Jordan KP, Jordan JL, Bedson J, Dziedzic KS. Quality indicators for the primary care of osteoarthritis: a systematic review. *Ann Rheum Dis* 2015;**74**:490–8. https://doi.org/10.1136/annrheumdis-2013-203913

146. Østerås N, Garratt A, Grotle M, Natvig B, Kjeken I, Kvien TK, Hagen KB. Patient-reported quality of care for osteoarthritis: development and testing of the osteoarthritis quality indicator questionnaire. *Arthritis Care Res* 2013;**65**:1043–51. https://doi.org/10.1002/acr.21976

147. Finney A. *Multisite peripheral joint pain: prevalence, impact and multidisciplinary support in community dwelling older adults*. Unpublished PhD thesis. Keele: Keele University; 2014.

148. Edwards J, Paskins Z, Hassell A. *Hands On: The Approach to the Patients Presenting with Multiple Joint Pain*. 2012. URL: www.arthritisresearchuk.org/health-professionals-and-students/reports/hands-on/hands-on-autumn-2012.aspx (accessed 11 May 2017).

149. Jordan KP, Kadam UT, Hayward R, Porcheret M, Young C, Croft P. Annual consultation prevalence of regional musculoskeletal problems in primary care: an observational study. *BMC Musculoskelet Disord* 2010;**11**:144. https://doi.org/10.1186/1471-2474-11-144

150. Croft P. The question is not "have you got it"? But "how much of it have you got"? *Pain* 2009;**141**:6–7. https://doi.org/10.1016/j.pain.2008.10.019

151. Schmidt CO, Baumeister SE. Simple patterns behind complex spatial pain reporting? Assessing a classification of multisite pain reporting in the general population. *Pain* 2007;**133**:174–82. https://doi.org/10.1016/j.pain.2007.04.022

152. Jordan H, Roderick P, Martin D. The Index of Multiple Deprivation 2000 and accessibility effects on health. *J Epidemiol Community Health* 2004;**58**:250–7. https://doi.org/10.1136/jech.2003.013011

153. National Institute for Health and Care Excellence. *Osteoarthritis: Care and Management in Adults*. *Clinical Guideline CG177*. London: NICE; 2014. URL: www.nice.org.uk/CG177 (accessed 11 May 2017).

154. Smith R. *Self-reported physical activity levels: measurements and assessment in community dwelling adults with, or at risk of osteoarthritis*. Unpublished PhD thesis. Keele: Keele University; 2017.

155. Department of Health. *At Least Five a Week. Evidence on the Impact of Physical Activity and its Relationship to Health. A report from the Chief Medical Officer*. London: DH; 2004.

156. Finney A, Healey E, Jordan Jl, Ryan S, Dziedzic KS. Multidisciplinary approaches to managing osteoarthritis in multiple joint sites: a systematic review. *BMC Musculoskel Disord* 2016;**17**:266. https://doi.org/10.1186/s12891-016-1125-5

157. Campbell M, Fitzpatrick R, Haines A, Kinmonth AL, Sandercock P, Spiegelhalter D, Tyrer P. Framework for design and evaluation of complex interventions to improve health. *BMJ* 2000;**321**:694–6. https://doi.org/10.1136/bmj.321.7262.694

158. The Cochrane Collaboration. *Effective Practice and Organisation of Care Group. Reviews*. URL: http://epoc.cochrane.org/our-reviews (accessed 19 April 2017).

159. Porcheret M, Main C, Croft P, Dziedzic K. Enhancing delivery of osteoarthritis care in the general practice consultation: evaluation of a behaviour change intervention. *BMC Fam Pract* 2018;**19**:26. https://doi.org/10.1186/s12875-018-0715-8

160. Healey EL, Main CJ, Ryan S, McHugh GA, Porcheret M, Finney AG, Morden A, Dziedzic KS. A nurse-led clinic for patients consulting with osteoarthritis in general practice: development and impact of training in a cluster randomised controlled trial. *BMC Fam Pract* 2016;**17**:173. https://doi.org/10.1186/s12875-016-0568-y

161. Smucker DR, Konrad TR, Curtis P, Carey TS. Practitioner self-confidence and patient outcomes in acute low back pain. *Arch Fam Med* 1998;**7**:223–8. https://doi.org/10.1001/archfami.7.3.223

162. Kurtz S, Silverman J, Benson J, Draper J. Marrying content and process in clinical method teaching: enhancing the Calgary-Cambridge guides. *Acad Med* 2003;**78**:802–9. https://doi.org/10.1097/00001888-200308000-00011

163. Research Institute of Primary Care and Health Sciences. *A Guide for People Who Have osteoarthritis*. Keele University. URL: www.keele.ac.uk/media/keeleuniversity/ri/primarycare/pdfs/OA_Guidebook.pdf/ (accessed 4 July 2017).

164. Arthritis Research UK. *Hands On Approach to Osteoarthritis*. URL: www.arthritisresearchuk.org/arthritis-information/arthritis-today-magazine/153-summer-2011/hands-on-approach.aspx (accessed 11 April 2017).

165. Edwards JJ. *Quality Indicators for the Care of Osteoarthritis in General Practice: Identification, Synthesis and Implementation*. PhD thesis. Keele: Keele University; 2017.

166. Blackburn S, Higginbottom A, Taylor R, Bird J, Østeräs N, Hagen KB, *et al.* Patient-reported quality indicatros for osteoarthritis: a patient and public generated self-report measure for primary care. *Res Involve Engage* 2016;**2**:5. https://doi.org/10.1186/s40900-016-0019-x

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

167. Jordan K, Edwards J, Porcheret M, Healey E, Jinks C, Bedson J, *et al.* Effect of a model consultation informed by guidelines on recorded quality of care of osteoarthritis (MOSAICS): a cluster randomised controlled trial in primary care. *Osteoarthritis Cartilage* 2017; in press. https://doi.org/10.1016/j.joca.2017.05.017

168. Dziedzic KS, Healey EL, Porcheret M, Afolabi EK, Lewis M, Morden A, *et al.* Implementing core nice guidelines for osteoarthritis in primary care with a model consultation (MOSAICS): a cluster randomised controlled trial. *Osteoarthritis Cartilage* 2018;**26**:43–53. https://doi.org/10.1016/j.joca.2017.09.010

169. Jinks C, Lewis M, Ong BN, Croft P. A brief screening tool for knee pain in primary care. 1. Validity and reliability. *Rheumatology* 2001;**40**:528–36. https://doi.org/10.1093/rheumatology/40.5.528

170. Washburn RA, Smith KW, Jette AM, Janney CA. The Physical Activity Scale for the Elderly (PASE): development and evaluation. *J Clin Epidemiol* 1993;**46**:153–62. https://doi.org/10.1016/0895-4356(93)90053-4

171. Ware J, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Med Care* 1996;**34**:220–33. https://doi.org/10.1097/00005650-199603000-00003

172. Oppong R, Jowett S, Lewis M, Clarkson K, Paskins Z, Croft P, *et al.* Cost-effectiveness of a model consultation to support self-management in patients with osteoarthritis. *Rheumatology* Oxford 2018;key:037. https://doi.org/10.1093/rheumatology/key037

173. Personal Social Services Research Unit. *Unit Costs of Health and Social Care 2012*. Canterbury: University of Kent; 2012.

174. Joint Formulary Committee. *British National Formulary*. 67 edn. London: Pharmaceutical Press; 2013.

175. Department of Health. *NHS Reference Costs 2012/13*. URL: www.gov.uk/government/publications/nhs-reference-costs-2012-to-2013 (accessed 11 May 2017).

176. Morden A, Jinks C, Ong BN, Porcheret M, Dziedzic KS. Acceptability of a 'guidebook' for the management of Osteoarthritis: a qualitative study of patient and clinician's perspectives. *BMC Musculoskelet Disord* 2014;**15**:427. https://doi.org/10.1186/1471-2474-15-427

177. Morden A, Brooks L, Jinks C, Porcheret M, Ong BN, Dziedzic K. Research 'push', long term-change, and general practice. *J Health Organ Manag* 2015;**29**:798–821. https://doi.org/10.1108/JHOM-07-2014-0119

178. Ong BN, Morden A, Brooks L, Porcheret M, Edwards JJ, Sanders T, *et al.* Changing policy and practice: making sense of national guidelines for osteoarthritis. *Soc Sci Med* 2014;**106**:101–9. https://doi.org/10.1016/j.socscimed.2014.01.036

179. Mays N, Pope C. Qualitative research: Observational methods in health care settings. *BMJ* 1995;**311**:182–4. https://doi.org/10.1136/bmj.311.6998.182

180. Morden A, Jinks C, Ong BN. Understanding help seeking for chronic joint pain: implications for providing supported self-management. *Qual Health Res* 2014;**24**:957–68. https://doi.org/10.1177/1049732314539853

181. Ong BN, Rogers A, Kennedy A, Bower P, Sanders T, Morden A, *et al.* Behaviour change and social blinkers? The role of sociology in trials of self-management behaviour in chronic conditions. *Sociol Health Illn* 2014;**36**:226–38. https://doi.org/10.1111/1467-9566.12113

182. May T. *Social Research: Issues, Method and Process*. Maidenhead: Open University Press; 2002.

183. Bowling A. *Research Methods in Health: Investigating Health and Health Services*. Maidenhead: Open University Press; 2001.

184. Charmaz K. *Constructing Grounded Theory: A Practical Guide Through Qualitative Analysis*. Thousand Oaks, CA: Sage; 2006.

185. Sharkey S, Larsen JA. Ethnographic Exploration: Participation and Meaning in Everyday Life. In Holloway I, editor. *Qualitative Methods in Health Research*. Maidenhead: Open University Press; 2005. pp. 168–90.

186. Dziedzic KS, Healey EL, Main CJ. Implementing the NICE osteoarthritis guidelines in primary care: a role for practice nurses. *Musculoskeletal Care* 2013;**11**:1–2. https://doi.org/10.1002/msc.1040

187. Brierley G, Brabyn S, Torgerson D, Watson J. Bias in recruitment to cluster randomized trials: a review of recent publications. *J Eval Clin Pract* 2012;**18**:878–86. https://doi.org/10.1111/j.1365-2753.2011.01700.x

188. Foster NE, Healey EL, Holden MA, Nicholls E, Whitehurst DG, Jowett S, *et al.* A multicentre, pragmatic, parallel group, randomised controlled trial to compare the clinical and cost-effectiveness of three physiotherapy-led exercise interventions for knee osteoarthritis in older adults: the BEEP trial protocol (ISRCTN: 93634563). *BMC Musculoskelet Disord* 2014;**15**:254. https://doi.org/10.1186/1471-2474-15-254

189. Peat G, McCarney R, Croft P. Knee pain and osteoarthritis in older adults: a review of community burden and current use of primary health care. *Ann Rheum Dis* 2001;**60**:91–7. https://doi.org/10.1136/ard.60.2.91

190. Duncan RC, Hay EM, Saklatvala J, Croft PR. Prevalence of radiographic osteoarthritis – it all depends on your point of view. *Rheumatology* 2006;**45**:757–60. https://doi.org/10.1093/rheumatology/kei270

191. Cibere J. Do we need radiographs to diagnose osteoarthritis? *Best Pract Res Clin Rheumatol* 2006;**20**:27–38. https://doi.org/10.1016/j.berh.2005.08.001

192. Juhl C, Christensen R, Roos EM, Zhang W, Lund H. Impact of exercise type and dose on pain and disability in knee osteoarthritis: a systematic review and meta-regression analysis of randomized controlled trials. *Arthritis Rheumatol* 2014;**66**:622–36. https://doi.org/10.1002/art.38290

193. Smidt N, de Vet HC, Bouter LM, Dekker J, Arendzen JH, de Bie RA, *et al.* Effectiveness of exercise therapy: a best-evidence summary of systematic reviews. *Aust J Physiother* 2005;**51**:71–85. https://doi.org/10.1016/S0004-9514(05)70036-2

194. Taylor NF, Dodd KJ, Shields N, Bruder A. Therapeutic exercise in physiotherapy practice is beneficial: a summary of systematic reviews 2002–5. *Aust J Physiother* 2007;**53**:7–16. https://doi.org/10.1016/S0004-9514(07)70057-0

195. Hurley MV. Muscle dysfunction and effective rehabilitation of knee osteoarthritis: what we know and what we need to find out. *Arthritis Rheum* 2003;**49**:444–52. https://doi.org/10.1002/art.11053

196. van Baar ME, Assendelft WJ, Dekker J, Oostendorp RA, Bijlsma JW. Effectiveness of exercise therapy in patients with osteoarthritis of the hip or knee: a systematic review of randomized clinical trials. *Arthritis Rheum* 1999;**42**:1361–9. https://doi.org/10.1002/1529-0131(199907)42:7<1361::AID-ANR9>3.0.CO;2-9

197. Hurley M, Dziedzic K, Bearne L, Sim J, Bury T. *The Clinical and Cost Effectiveness of Physiotherapy in the Managment of Elderly People with Common Rheumatological Conditions: Evidence Briefing*. London: Chartered Society of Physiotherapy; 2002.

198. Puett DW, Griffin MR. Published trials of nonmedicinal and noninvasive therapies for hip and knee osteoarthritis. *Ann Intern Med* 1994;**121**:133–40. https://doi.org/10.7326/0003-4819-121-2-199407150-00010

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

199. Holden MA, Nicholls EE, Hay EM, Foster NE. Physical therapists' use of therapeutic exercise for patients with clinical knee osteoarthritis in the United kingdom: in line with current recommendations? *Phys Ther* 2008;**88**:1109–21. https://doi.org/10.2522/ptj.20080077

200. Pham T, Van Der Heijde D, Lassere M, Altman RD, Anderson JJ, Bellamy N, *et al.* Outcome variables for osteoarthritis clinical trials: The OMERACT-OARSI set of responder criteria. *J Rheumatol* 2003;**30**:1648–54.

201. Pham T, van der Heijde D, Altman RD, Anderson JJ, Bellamy N, Hochberg M, *et al.* OMERACT-OARSI initiative: Osteoarthritis Research Society International set of responder criteria for osteoarthritis clinical trials revisited. *Osteoarthr Cartil* 2004;**12**:389–99. https://doi.org/10.1016/j.joca.2004.02.001

202. Hay E, Barlas P, Foster N, Hill J, Thomas E, Young J. Is acupuncture a useful adjunct to physiotherapy for older adults with knee pain?: the "acupuncture, physiotherapy and exercise" (APEX) study [ISRCTN88597683.] *BMC Musculoskelet Disord* 2004;**5**:31. https://doi.org/10.1186/1471-2474-5-31

203. Foster NE, Thomas E, Barlas P, Hill JC, Young J, Mason E, Hay EM. Acupuncture as an adjunct to exercise based physiotherapy for osteoarthritis of the knee: randomised controlled trial. *BMJ* 2007;**335**:436. https://doi.org/10.1136/bmj.39280.509803.BE

204. Pisters MF, Veenhof C, van Meeteren NL, Ostelo RW, de Bakker DH, Schellevis FG, Dekker J. Long-term effectiveness of exercise therapy in patients with osteoarthritis of the hip or knee: a systematic review. *Arthritis Rheum* 2007;**57**:1245–53. https://doi.org/10.1002/art.23009

205. Marks R, Allegrante JP. Chronic osteoarthritis and adherence to exercise: a review of the literature. *J Aging Phys Act* 2005;**13**:434–60. https://doi.org/10.1123/japa.13.4.434

206. Hillsdon M, Foster C, Thorogood M. Interventions for promoting physical activity. *Cochrane Database Syst Rev* 2005;**1**:CD003180.

207. Rankin G, Rushton A, Olver P, Moore A. Chartered Society of Physiotherapy's identification of national research priorities for physiotherapy using a modified Delphi technique. *Physiotherapy* 2012;**98**:260–72. https://doi.org/10.1016/j.physio.2012.03.002

208. Jordan JL, Holden MA, Mason EE, Foster NE. Interventions to improve adherence to exercise for chronic musculoskeletal pain in adults. *Cochrane Database Syst Rev* 2010;**1**:CD005956. https://doi.org/10.1002/14651858.CD005956.pub2

209. Bennell KL, Hinman RS. A review of the clinical evidence for exercise in osteoarthritis of the hip and knee. *J Sci Med Sport* 2011;**14**:4–9. https://doi.org/10.1016/j.jsams.2010.08.002

210. Marks R. Knee osteoarthritis and exercise adherence: a review. *Curr Aging Sci* 2012;**5**:72–83. https://doi.org/10.2174/1874609811205010072

211. World Health Organization. *Adherence to Long-term Therapies: Evidence for Action*. Geneva: World Health Organization; 2003.

212. Deyo RA. Compliance with therapeutic regimens in arthritis: issues, current status, and a future agenda. *Semin Arthritis Rheum* 1982;**12**:233–44. https://doi.org/10.1016/0049-0172(82)90063-4

213. Bird SR, Smith A, James K. *Exercise Benefits and Prescriptions*. Devon: Stanley Thornes Ltd; 1998.

214. Bandura A. *Social Foundations of Thought and Action: A Social Cognitive Theory*. Englewood Cliffs, NJ: Prentice Hall; 1986.

215. McAuley E, Jerome GJ, Marquez DX, Elavsky S, Blissmer B. Exercise self-efficacy in older adults: social, affective, and behavioral influences. *Ann Behav Med* 2003;**25**:1–7. https://doi.org/10.1207/S15324796ABM2501_01

216. Holden MA, Nicholls EE, Young J, Hay EM, Foster NE. Role of exercise for knee pain: what do older adults in the community think? *Arthritis Care Res* 2012;**64**:1554–64. https://doi.org/10.1002/acr.21700

217. Mutrie N, Doolin O, Fitzsimons CF, Grant PM, Granat M, Grealy M, *et al.* Increasing older adults' walking through primary care: results of a pilot randomized controlled trial. *Fam Pract* 2012;**29**:633–42. https://doi.org/10.1093/fampra/cms038

218. Bravata DM, Smith-Spangler C, Sundaram V, Gienger AL, Lin N, Lewis R, *et al.* Using pedometers to increase physical activity and improve health: a systematic review. *JAMA* 2007;**298**:2296–304. https://doi.org/10.1001/jama.298.19.2296

219. Cameron LD, Leventhal H. *The Self-regulation of Health and Illness Behaviour*. London: Routledge; 2003.

220. Holden MA, Nicholls EE, Young J, Hay EM, Foster NE. UK-based physical therapists' attitudes and beliefs regarding exercise and knee osteoarthritis: findings from a mixed-methods study. *Arthritis Rheum* 2009;**61**:1511–21. https://doi.org/10.1002/art.24829

221. Partridge MR. Translating research into practice: how are guidelines implemented? *Eur Respir J Suppl* 2003;**39**:23s–29s. https://doi.org/10.1183/09031936.0.00054503

222. Peter WF, van der Wees PJ, Verhoef J, de Jong Z, van Bodegom-Vos L, Hilberdink WK, *et al.* Postgraduate education to increase adherence to a Dutch physiotherapy practice guideline for hip and knee OA: a randomized controlled trial. *Rheumatology* 2013;**52**:368–75. https://doi.org/10.1093/rheumatology/kes264

223. Rebbeck T, Macedo LG, Maher CG. Compliance with clinical guidelines for whiplash improved with a targeted implementation strategy: a prospective cohort study. *BMC Health Serv Res* 2013;**13**:213. https://doi.org/10.1186/1472-6963-13-213

224. Scott SD, Albrecht L, O'Leary K, Ball GDC, Hartling L, Hofmeyer A, *et al.* Systematic review of knowledge translation strategies in the allied health professions. *Implement Sci* 2012;**7**:70. https://doi.org/10.1186/1748-5908-7-70

225. Cherry MG, Brown JM, Bethell GS, Neal T, Shaw NJ. Features of educational interventions that lead to compliance with hand hygiene in healthcare professionals within a hospital care setting. A BEME systematic review: BEME Guide No. 22. *Med Teach* 2012;**34**:e406–20. https://doi.org/10.3109/0142159X.2012.680936

226. Grol R, Grimshaw J. From best evidence to best practice: effective implementation of change in patients' care. *Lancet* 2003;**362**:1225–30. https://doi.org/10.1016/S0140-6736(03)14546-1

227. van der Wees PJ, Jamtvedt G, Rebbeck T, de Bie RA, Dekker J, Hendriks EJ. Multifaceted strategies may increase implementation of physiotherapy clinical guidelines: a systematic review. *Aust J Physiother* 2008;**54**:233–41. https://doi.org/10.1016/S0004-9514(08)70002-3

228. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. *J Rheumatol* 1988;**15**:1833–40.

229. Von Korff M, Ormel J, Keefe FJ, Dworkin SF. Grading the severity of chronic pain. *Pain* 1992;**50**:133–49. https://doi.org/10.1016/0304-3959(92)90154-4

230. Dunn KM, Croft PR, Main CJ, Von Korff M. A prognostic approach to defining chronic pain: replication in a UK primary care low back pain population. *Pain* 2008;**135**:48–54. https://doi.org/10.1016/j.pain.2007.05.001

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

231. Gov.uk. *Health and Social Care Information Centre. Exeter System*. URL: www.gov.uk/government/organisations/health-and-social-care-information-centre (accessed 2011).

232. O'Reilly SC, Muir KR, Doherty M. Screening for pain in knee osteoarthritis: which question? *Ann Rheum Dis* 1996;**55**:931–3. https://doi.org/10.1136/ard.55.12.931

233. Trudeau J, Van Inwegen R, Eaton T, Bhat G, Paillard F, Ng D, *et al.* Assessment of Pain and activity using an electronic pain diary and actigraphy device in a randomized, placebo-controlled crossover trial of celecoxib in osteoarthritis of the knee. *Pain Pract* 2015;**15**:247–55. https://doi.org/10.1111/papr.12167

234. Borkovec TD, Nau SD. Credibility of analogue therapy rationales. *J Behav Ther Exp Psy* 1972;**3**:257–60. https://doi.org/10.1016/0005-7916(72)90045-6

235. Devilly GJ, Borkovec TD. Psychometric properties of the credibility/expectancy questionnaire. *J Behav Ther Exp Psychiatry* 2000;**31**:73–86. https://doi.org/10.1016/S0005-7916(00)00012-4

236. Broadbent E, Petrie KJ, Main J, Weinman J. The brief illness perception questionnaire. *J Psychosom Res* 2006;**60**:631–7. https://doi.org/10.1016/j.jpsychores.2005.10.020

237. Resnick B, Luisi D, Vogel A, Junaleepa P. Reliability and validity of the self-efficacy for exercise and outcome expectations for exercise scales with minority older adults. *J Nurs Meas* 2004;**12**:235–47. https://doi.org/10.1891/jnum.12.3.235

238. Resnick B. Reliability and validity of the Outcome Expectations for Exercise Scale-2. *J Aging Phys Act* 2005;**13**:382–94. https://doi.org/10.1123/japa.13.4.382

239. Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med* 2006;**166**:1092–7. https://doi.org/10.1001/archinte.166.10.1092

240. Kroenke K, Strine TW, Spitzer RL, Williams JB, Berry JT, Mokdad AH. The PHQ-8 as a measure of current depression in the general population. *J Affect Disord* 2009;**114**:163–73. https://doi.org/10.1016/j.jad.2008.06.026

241. EuroQol G. EuroQol – a new facility for the measurement of health-related quality of life. *Health Policy* 1990;**16**:199–208. https://doi.org/10.1016/0168-8510(90)90421-9

242. Department of Health. *Start Active, Stay Active: A Report on Physical Activity for Health from the Four Home Countries' Chief Medical Officers*. London: DH; 2011.

243. Terwee CB, Roorda LD, Dekker J, Bierma-Zeinstra SM, Peat G, Jordan KP, *et al.* Mind the MIC: large variation among populations and methods. *J Clin Epidemiol* 2010;**63**:524–34. https://doi.org/10.1016/j.jclinepi.2009.08.010

244. Roberts C, Roberts SA. Design and analysis of clinical trials with clustering effects due to treatment. *Clin Trials* 2005;**2**:152–62. https://doi.org/10.1191/1740774505cn076oa

245. Boutron I, Moher D, Altman DG, Schulz KF, Ravaud P, CONSORT Group. Extending the CONSORT statement to randomized trials of nonpharmacologic treatment: explanation and elaboration. *Ann Intern Med* 2008;**148**:295–309. https://doi.org/10.7326/0003-4819-148-4-200802190-00008

246. Boutron I, Moher D, Altman DG, Schulz KF, Ravaud P, CONSORT Group. Methods and processes of the CONSORT Group: example of an extension for trials assessing nonpharmacologic treatments. *Ann Intern Med* 2008;**148**:W60–6. https://doi.org/10.7326/0003-4819-148-4-200802190-00008-w1

247. Zwarenstein M, Treweek S, Gagnier JJ, Altman DG, Tunis S, Haynes B, *et al.* Improving the reporting of pragmatic trials: an extension of the CONSORT statement. *BMJ* 2008;**337**:a2390. https://doi.org/10.1136/bmj.a2390

248. Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux PJ, *et al.* CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *Int J Surg* 2012;**10**:28–55. https://doi.org/10.1016/j.ijsu.2011.10.001

249. Peduzzi P, Henderson W, Hartigan P, Lavori P. Analysis of randomized controlled trials. *Epidemiol Rev* 2002;**24**:26–38. https://doi.org/10.1093/epirev/24.1.26

250. ActiGraph Support. *How Does Bout Detection Work and Where Did We Get Our Defaults?* 2012. URL: https://help.theactigraph.com/entries/21684552 (accessed 24 March 2014).

251. Song J, Semanik P, Sharma L, Chang RW, Hochberg MC, Mysiw WJ, *et al.* Assessing physical activity in persons with knee osteoarthritis using accelerometers: data from the osteoarthritis initiative. *Arthritis Care Res* 2010;**62**:1724–32. https://doi.org/10.1002/acr.20305

252. Freedson PS, Melanson E, Sirard J. Calibration of the Computer Science and Applications, Inc. accelerometer. *Med Sci Sports Exerc* 1998;**30**:777–81. https://doi.org/10.1097/00005768-199805000-00021

253. Dolan P. Modeling valuations for EuroQol health states. *Med Care* 1997;**35**:1095–108. https://doi.org/10.1097/00005650-199711000-00002

254. Morris S, Devlin N, Parkin D. *Economic Analysis in Health Care*. Chichester: John Wiley & Sons; 2007.

255. Briggs A, Fenn P. Confidence intervals or surfaces? Uncertainty on the cost-effectiveness plane. *Health Econ* 1998;**7**:723–40. https://doi.org/10.1002/(SICI)1099-1050(199812)7:8<723::AID-HEC392>3.0.CO;2-O

256. Fenwick E, Byford S. A guide to cost-effectiveness acceptability curves. *Br J Psychiatry* 2005;**187**:106–8. https://doi.org/10.1192/bjp.187.2.106

257. Manca A, Hawkins N, Sculpher MJ. Estimating mean QALYs in trial-based cost-effectiveness analysis: the importance of controlling for baseline utility. *Health Econ* 2005;**14**:487–96. https://doi.org/10.1002/hec.944

258. Office for National Statistics. *Standard Occupational Classification: SOC2010 Volume 2 The Structure and Index*. London: Office for National Statistics; 2010.

259. Office for National Statistics. *Annual Survey of Hours and Earnings*. London: Office for National Statistics; 2013.

260. Lewin S, Glenton C, Oxman AD. Use of qualitative methods alongside randomised controlled trials of complex healthcare interventions: methodological study. *BMJ* 2009;**339**:b3496. https://doi.org/10.1136/bmj.b3496

261. O'Cathain A, Thomas KJ, Drabble SJ, Rudolph A, Hewison J. What can qualitative research do for randomised controlled trials? A systematic mapping review. *BMJ Open* 2013;**3**:e002889. https://doi.org/10.1136/bmjopen-2013-002889

262. Plano Clark VL, Schumacher K, West CM, Edrington J, *et al.* Practices for embedding an interpretive qualitative approach within a randomized clinical trial. *J Mix Method Res* 2013;**7**:219–42. https://doi.org/10.1177/1558689812474372

263. Saldana J. *Longitudinal Qualitative Research. Analysis Change Through Time*. Walnut Creek, CA: AltaMira Press; 2003.

264. MacFarlane GJ, Croft PR, Schollum J, Silman AJ. Widespread pain: is an improved classification possible? *J Rheumatol* 1996;**23**:1628–32.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

265. Kigozi J, Jowett S, Nicholls E, Tooth S, Tooth S, Hay EM, Foster NE*et al.* Cost-utility analysis of interventions to improve effectiveness of exercise therapy for adults with knee osteoarthritis: the BEEP trial [published online ahead of print June 6 2018]. *Rheumatol Adv Pract* 2018. https://doi.org/10.1093/rap/rky018

266. Moore A, Holden M, Foster N, Jinks C. *Uptake and maintenance of physiotherapy-led exercise and general physical activity in older adults with knee pain: the BEEP longitudinal qualitative study*. Keele: Keele University; 2017.

267. Hall AM, Ferreira PH, Maher CG, Latimer J, Ferreira ML. The influence of the therapist-patient relationship on treatment outcome in physical rehabilitation: a systematic review. *Phys Ther* 2010;**90**:1099–110. https://doi.org/10.2522/ptj.20090245

268. Petursdottir U, Arnadottir SA, Halldorsdottir S. Facilitators and barriers to exercising among people with osteoarthritis: a phenomenological study. *Phys Ther* 2010;**90**:1014–25. https://doi.org/10.2522/ptj.20090217

269. Pinto D, Robertson MC, Hansen P, Abbott JH. Cost-effectiveness of nonpharmacologic, nonsurgical interventions for hip and/or knee osteoarthritis: systematic review. *Value Health* 2012;**15**:1–12. https://doi.org/10.1016/j.jval.2011.09.003

270. Hurley MV, Walsh NE, Mitchell H, Nicholas J, Patel A. Long-term outcomes and costs of an integrated rehabilitation program for chronic knee pain: a pragmatic, cluster randomized, controlled trial. *Arthritis Care Res* 2012;**64**:238–47. https://doi.org/10.1002/acr.20642

271. Bennell KL, Kyriakides M, Metcalf B, Egerton T, Wrigley TV, Hodges PW, *et al.* Neuromuscular versus quadriceps strengthening exercise in patients with medial knee osteoarthritis and varus malalignment: a randomized controlled trial. *Arthritis Rheumatol* 2014;**66**:950–9. https://doi.org/10.1002/art.38317

272. Øiestad BE, Østerås N, Frobell R, Grotle M, Brøgger H, Risberg MA. Efficacy of strength and aerobic exercise on patient-reported outcomes and structural changes in patients with knee osteoarthritis: study protocol for a randomized controlled trial. *BMC Musculoskelet Disord* 2013;**14**:266. https://doi.org/10.1186/1471-2474-14-266

273. Messier SP, Mihalko SL, Beavers DP, Nicklas BJ, DeVita P, Carr JJ, *et al.* Strength Training for Arthritis Trial (START): design and rationale. *BMC Musculoskelet Disord* 2013;**14**:208. https://doi.org/10.1186/1471-2474-14-208

274. Isaacs AJ, Critchley JA, Tai SS, Buckingham K, Westley D, Harridge SD, *et al.* Exercise Evaluation Randomised Trial (EXERT): a randomised trial comparing GP referral for leisure centre-based exercise, community-based walking and advice only. *Health Technol Assess* 2007;**11**(10). https://doi.org/10.3310/hta11100

275. Salisbury C, Foster NE, Hopper C, Bishop A, Hollinghurst S, Coast J, *et al.* A pragmatic randomised controlled trial of the effectiveness and cost-effectiveness of 'PhysioDirect' telephone assessment and advice services for physiotherapy. *Health Technol Assess* 2013;**17**(2). https://doi.org/10.3310/hta17020

276. Deyle GD, Gill NW, Allison SC, Hando BR, Rochino DA. Knee OA: which patients are unlikely to benefit from manual PT and exercise? *J Fam Pract* 2012;**61**:E1–8.

277. Arthritis Research UK. *Osteoarthritis: An Information Booklet*. 2004. URL: www.arthritisresearchuk.org/arthritis-information/conditions/osteoarthritis.aspx (accessed 11 May 2017).

278. Porcheret M, Jordan K, Croft P, Primary Care Rhumatology Society. Treatment of knee pain in older adults in primary care: development of an evidence-based model of care. *Rheumatology* 2007;**46**:638–48. https://doi.org/10.1093/rheumatology/kel340

279. Lin EH, Tang L, Katon W, Hegel MT, Sullivan MD, Unützer J. Arthritis pain and disability: response to collaborative depression care. *Gen Hosp Psychiatry* 2006;**28**:482–6. https://doi.org/10.1016/j.genhosppsych.2006.08.006

280. Freeling P, Rao BM, Paykel ES, Sireling LI, Burton RH. Unrecognised depression in general practice. *Br Med J* 1985;**290**:1880–3. https://doi.org/10.1136/bmj.290.6485.1880

281. Arroll B, Khin N, Kerse N. Screening for depression in primary care with two verbally asked questions: cross sectional study. *BMJ* 2003;**327**:1144–6. https://doi.org/10.1136/bmj.327.7424.1144

282. National Institute for Health and Care Excellence. *NICE Quality and Outcomes Framework Indicator. Depression*. URL: www.nice.org.uk/standards-and-indicators/qofindicators?categories=3895&page=1 (accessed 19 April 2017).

283. Katon W, Unützer J, Fan MY, Williams JW, Schoenbaum M, Lin EH, Hunkeler EM. Cost-effectiveness and net benefit of enhanced treatment of depression for older adults with diabetes and depression. *Diabetes Care* 2006;**29**:265–70. https://doi.org/10.2337/diacare.29.02.06.dc05-1572

284. Gilbody S, House AO, Sheldon TA. Screening and case finding instruments for depression. *Cochrane Database Syst Rev* 2005;**4**:CD002792. https://doi.org/10.1002/14651858.cd002792.pub2

285. Gilbody S, Sheldon T, Wessely S. Should we screen for depression? *BMJ* 2006;**332**:1027–30. https://doi.org/10.1136/bmj.332.7548.1027

286. Katon W, Lin EH, Kroenke K. The association of depression and anxiety with medical symptom burden in patients with chronic medical illness. *Gen Hosp Psychiatry* 2007;**29**:147–55. https://doi.org/10.1016/j.genhosppsych.2006.11.005

287. National Institute for Health and Care Excellence. *The Treatment and Management of Depression in Adults with Chronic Physical Health Problems*. *Clinical Guideline CG91*. London: NICE; 2009.

288. Mitchell AJ, Coyne JC. Do ultra-short screening instruments accurately detect depression in primary care? A pooled analysis and meta-analysis of 22 studies. *Br J Gen Pract* 2007;**57**:144–51.

289. Pignone MP, Gaynes BN, Rushton JL, Burchell CM, Orleans CT, Mulrow CD, Lohr KN. Screening for depression in adults: a summary of the evidence for the U.S. Preventive Services Task Force. *Ann Intern Med* 2002;**136**:765–76. https://doi.org/10.7326/0003-4819-136-10-200205210-00013

290. Nisenson LG, Pepper CM, Schwenk TL, Coyne JC. The nature and prevalence of anxiety disorders in primary care. *Gen Hosp Psychiatry* 1998;**20**:21–8. https://doi.org/10.1016/S0163-8343(97)00096-0

291. Kroenke K, Spitzer RL, Williams JB, Monahan PO, Löwe B. Anxiety disorders in primary care: prevalence, impairment, comorbidity, and detection. *Ann Intern Med* 2007;**146**:317–25. https://doi.org/10.7326/0003-4819-146-5-200703060-00004

292. Memel DS, Kirwan JR, Sharp DJ, Hehir M. General practitioners miss disability and anxiety as well as depression in their patients with osteoarthritis. *Br J Gen Pract* 2000;**50**:645–8.

293. Addolorato G, Mirijello A, D'Angelo C, Leggio L, Ferrulli A, Abenavoli L, *et al.* State and trait anxiety and depression in patients affected by gastrointestinal diseases: psychometric evaluation of 1641 patients referred to an internal medicine outpatient setting. *Int J Clin Pract* 2008;**62**:1063–9. https://doi.org/10.1111/j.1742-1241.2008.01763.x

294. Jackson JL, Passamonti M, Kroenke K. Outcome and impact of mental disorders in primary care at 5 years. *Psychosom Med* 2007;**69**:270–6. https://doi.org/10.1097/PSY.0b013e3180314b59

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

295. Creamer P, Lethbridge-Cejku M, Costa P, Tobin JD, Herbst JH, Hochberg MC. The relationship of anxiety and depression with self-reported knee pain in the community: data from the Baltimore Longitudinal Study of Aging. *Arthritis Care Res* 1999;**12**:3–7. https://doi.org/10.1002/1529-0131(199902)12:1<3::AID-ART2>3.0.CO;2-K

296. Lang AJ, Stein MB. Screening for anxiety in primary care: why bother? *Gen Hosp Psychiatry* 2002;**24**:365–6. https://doi.org/10.1016/S0163-8343(02)00216-5

297. Katon W, Roy-Byrne P. Anxiety disorders: efficient screening is the first step in improving outcomes. *Ann Intern Med* 2007;**146**:390–2. https://doi.org/10.7326/0003-4819-146-5-200703060-00011

298. Eldridge SM, Ashby D, Feder GS. Informed patient consent to participation in cluster randomized trials: an empirical exploration of trials in primary care. *Clin Trials* 2005;**2**:91–8. https://doi.org/10.1191/1740774505cn070oa

299. Gilbody S, Sheldon T, House A. Screening and case-finding instruments for depression: a meta-analysis. *CMAJ* 2008;**178**:997–1003. https://doi.org/10.1503/cmaj.070281

300. Mallen CD, Nicholl BI, Lewis M, Bartlam B, Green D, Jowett S, *et al.* The effects of implementing a point-of-care electronic template to prompt routine anxiety and depression screening in patients consulting for osteoarthritis (the Primary Care Osteoarthritis Trial): a cluster randomised trial in primary care. *PLOS Med* 2017;**14**:e1002273.

301. Edwards SJ, Braunholtz DA, Lilford RJ, Stevens AJ. Ethical issues in the design and conduct of cluster randomised controlled trials. *BMJ* 1999;**318**:1407–9. https://doi.org/10.1136/bmj.318.7195.1407

302. Hutton JL. Are distinctive ethical principles required for cluster randomized controlled trials? *Stat Med* 2001;**20**:473–88. https://doi.org/10.1002/1097-0258(20010215)20:3<473::AID-SIM805>3.0.CO;2-D

303. National Institute for Health and Care Excellence. *Generalised Anxiety Disorder and panic Disorder (With or Without Agoraphobia) In Adults. Management in Primary, Secondary and Community Care*. *Clinical Guideline CG113*. London: NICE; 2011.

304. Pilling S, Anderson I, Goldberg D, Meader N, Taylor C, Two Guideline Development Groups. Depression in adults, including those with a chronic physical health problem: summary of NICE guidance. *BMJ* 2009;**339**:b4108. https://doi.org/10.1136/bmj.b4108

305. Sullivan MJL, Bishop SR, Pivik J. The pain catastrophizing scale: development and validation. *Psychol Assess*1995;**7**:524–32. https://doi.org/10.1037/1040-3590.7.4.524

306. Diggle PJ, Heagerty P, Liang K-Y, Zeger SL. *Analysis of Longitudinal Data*. 2nd ed. Oxford: Oxford University Press; 2002.

307. Carter BR, Hood K. Balance algorithm for cluster randomized trials. *BMC Med Res Methodol* 2008;**8**:65. https://doi.org/10.1186/1471-2288-8-65

308. Guest G, MacQueen KM, Namey EE. *Applied Thematic Analysis*. Thousand Oaks, CA: Sage; 2012. https://doi.org/10.4135/9781483384436

309. Creswell JW, Plano Clark VL. *Designing and Conducting Mixed Methods Research*. Thousand Oaks, CA: Sage; 2007.

310. Teddlie C, Yu F. Mixed methods sampling: a typology with examples. *J Mix Method Res* 2007;**1**:77–100. https://doi.org/10.1177/2345678906292430

311. Ritchie J, Spencer L. Qualitative Data Analysis for Applied Policy Research. In Bryman A, Burgess RG, editors. *Analysing Qualitative Data*. London: Routledge; 1994. pp. 173–94. https://doi.org/10.4324/9780203413081_chapter_9

312. Miles MB, Huberman AM. *Qualitative Data Analysis: An Expanded Sourcebook*. Thousand Oaks, CA: Sage; 1994.

313. Mitchell AJ. Screening for cancer-related distress: when is implementation successful and when is it unsuccessful? *Acta Oncol* 2013;**52**:216–24. https://doi.org/10.3109/0284186X.2012.745949

314. Maxwell M, Harris F, Hibberd C, Donaghy E, Pratt R, Williams C, *et al.* A qualitative study of primary care professionals' views of case finding for depression in patients with diabetes or coronary heart disease in the UK. *BMC Fam Pract* 2013;**14**:46. https://doi.org/10.1186/1471-2296-14-46

315. He Y, Zhang M, Lin EH, Bruffaerts R, Posada-Villa J, Angermeyer MC, *et al.* Mental disorders among persons with arthritis: results from the World Mental Health Surveys. *Psychol Med* 2008;**38**:1639–50. https://doi.org/10.1017/S0033291707002474

316. Mallen CD, Peat G, Thomas E, Dunn KM, Croft PR. Prognostic factors for musculoskeletal pain in primary care: a systematic review. *Br J Gen Pract* 2007;**57**:655–61.

317. Wright AA, Cook C, Abbott JH. Variables associated with the progression of hip osteoarthritis: a systematic review. *Arthritis Rheum* 2009;**61**:925–36. https://doi.org/10.1002/art.24641

318. Chapple CM, Nicholson H, Baxter GD, Abbott JH. Patient characteristics that predict progression of knee osteoarthritis: a systematic review of prognostic studies. *Arthritis Care Res* 2011;**63**:1115–25. https://doi.org/10.1002/acr.20492

319. Kwok WY, Plevier JW, Rosendaal FR, Huizinga TW, Kloppenburg M. Risk factors for progression in hand osteoarthritis: a systematic review. *Arthritis Care Res* 2013;**65**:552–62. https://doi.org/10.1002/acr.21851

320. Mallen CD, Peat G, Thomas E, Lacey R, Croft P. Predicting poor functional outcome in community-dwelling older adults with knee pain: prognostic value of generic indicators. *Ann Rheum Dis* 2007;**66**:1456–61. https://doi.org/10.1136/ard.2006.067975

321. Thomas E, Dunn KM, Mallen C, Peat G. A prognostic approach to defining chronic pain: application to knee pain in older adults. *Pain* 2008;**139**:389–97. https://doi.org/10.1016/j.pain.2008.05.010

322. Rosemann T, Gensichen J, Sauer N, Laux G, Szecsenyi J. The impact of concomitant depression on quality of life and health service utilisation in patients with osteoarthritis. *Rheumatol Int* 2007;**27**:859–63. https://doi.org/10.1007/s00296-007-0309-6

323. Scott KM, Von Korff M, Alonso J, Angermeyer MC, Bromet E, Fayyad J, *et al.* Mental-physical co-morbidity and its relationship with disability: results from the World Mental Health Surveys. *Psychol Med* 2009;**39**:33–43. https://doi.org/10.1017/S0033291708003188

324. Kroenke K, Wu J, Bair MJ, Krebs EE, Damush TM, Tu W. Reciprocal relationship between pain and depression: a 12-month longitudinal analysis in primary care. *J Pain* 2011;**12**:964–73. https://doi.org/10.1016/j.jpain.2011.03.003

325. Unützer J, Katon W, Callahan CM, Williams JW, Hunkeler E, Harpole L, *et al.* Collaborative care management of late-life depression in the primary care setting: a randomized controlled trial. *JAMA* 2002;**288**:2836–45. https://doi.org/10.1001/jama.288.22.2836

326. Roy-Byrne P, Craske MG, Sullivan G, Rose RD, Edlund MJ, Lang AJ, *et al.* Delivery of evidence-based treatment for multiple anxiety disorders in primary care: a randomized controlled trial. *JAMA* 2010;**303**:1921–8. https://doi.org/10.1001/jama.2010.608

327. Roy-Byrne P, Sullivan MD, Sherbourne CD, Golinelli D, Craske MG, Sullivan G, Stein MB. Effects of pain and prescription opioid use on outcomes in a collaborative care intervention for anxiety. *Clin J Pain* 2013;**29**:800–6. https://doi.org/10.1097/AJP.0b013e318278d475

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

328. Whooley MA, Avins AL, Miranda J, Browner WS. Case-finding instruments for depression. Two questions are as good as many. *J Gen Intern Med* 1997;**12**:439–45. https://doi.org/10.1046/j.1525-1497.1997.00076.x

329. Zimmermann C, Del Piccolo L, Finset A. Cues and concerns by patients in medical consultations: a literature review. *Psychol Bull* 2007;**133**:438–63. https://doi.org/10.1037/0033-2909.133.3.438

330. Del Piccolo L, Saltini A, Zimmermann C, Dunn G. Differences in verbal behaviours of patients with and without emotional distress during primary care consultations. *Psychol Med* 2000;**30**:629–43. https://doi.org/10.1017/S003329179900197X

331. Bensing JM, Verheul W, Jansen J, Langewitz WA. Looking for trouble: the added value of sequence analysis in finding evidence for the role of physicians in patients' disclosure of cues and concerns. *Med Care* 2010;**48**:583–8. https://doi.org/10.1097/MLR.0b013e3181d567a5

332. Hawker GA, Gignac MA, Badley E, Davis AM, French MR, Li Y, *et al.* A longitudinal study to explain the pain-depression link in older adults with osteoarthritis. *Arthritis Care Res* 2011;**63**:1382–90. https://doi.org/10.1002/acr.20298

333. Parmelee PA, Tighe CA, Dautovich ND. Sleep disturbance in osteoarthritis: linkages with pain, disability and depressive symptoms. *Arthritis Care Res* 2015;**67**:358–65. https://doi.org/10.1002/acr.22459

334. Bridle C, Spanjers K, Patel S, Atherton NM, Lamb SE. Effect of exercise on depression severity in older people: systematic review and meta-analysis of randomised controlled trials. *Br J Psychiatry* 2012;**201**:180–5. https://doi.org/10.1192/bjp.bp.111.095174

335. Cooney GM, Dwan K, Greig CA, Lawlor DA, Rimer J, Waugh FR, *et al.* Exercise for depression. *Cochrane Database Syst Rev* 2013;**9**:CD004366. https://doi.org/10.1002/14651858.CD004366.pub6

336. Maserejian NN, Fischer MA, Trachtenberg FL, Yu J, Marceau LD, McKinlay JB, Katz JN. Variations among primary care physicians in exercise advice, imaging, and analgesics for musculoskeletal pain: results from a factorial experiment. *Arthritis Care Res* 2014;**66**:147–56. https://doi.org/10.1002/acr.22143

337. Meijer A, Roseman M, Delisle VC, Milette K, Levis B, Syamchandra A, *et al.* Effects of screening for psychological distress on patient outcomes in cancer: a systematic review. *J Psychosom Res* 2013;**75**:1–17. https://doi.org/10.1016/j.jpsychores.2013.01.012

338. Thombs BD, Roseman M, Coyne JC, de Jonge P, Delisle VC, Arthurs E, *et al.* Does evidence support the American Heart Association's recommendation to screen patients for depression in cardiovascular care? An updated systematic review. *PLOS ONE* 2013;**8**:e52654. https://doi.org/10.1371/journal.pone.0052654

339. Thombs BD, Ziegelstein RC, Roseman M, Kloda LA, Ioannidis JP. There are no randomized controlled trials that support the United States Preventive Services Task Force Guideline on screening for depression in primary care: a systematic review. *BMC Med* 2014;**12**:13. https://doi.org/10.1186/1741-7015-12-13

340. Keshavarz H, Fitzpatrick-Lewis D, Streiner DL, Maureen R, Ali U, Shannon HS, Raina P. Screening for depression: a systematic review and meta-analysis. *CMAJ Open* 2013;**1**:E159–67. https://doi.org/10.9778/cmajo.20130030

341. Thombs BD. Routine depression screening for patients with diabetes. *JAMA* 2014;**312**:2412–13. https://doi.org/10.1001/jama.2014.14771

342. Thombs BD, Coyne JC. Moving forward by moving back: re-assessing guidelines for cancer distress screening. *J Psychosom Res* 2013;**75**:20–2. https://doi.org/10.1016/j.jpsychores.2013.05.002

343. Thombs BD, Ziegelstein RC. Evidence does matter – and evidence does not support the National Heart Foundation of Australia's consensus statement on depression screening. *J Psychosom Res* 2014;**76**:173–4. https://doi.org/10.1016/j.jpsychores.2013.12.005

344. Bland RC, Streiner DL. Why screening for depression in primary care is impractical. *CMAJ* 2013;**185**:753–4. https://doi.org/10.1503/cmaj.130634

345. Goldberg D. The value of screening in patient populations with high prevalence of a disorder. *BMC Med* 2014;**12**:14. https://doi.org/10.1186/1741-7015-12-14

346. Burton C, Simpson C, Anderson N. Diagnosis and treatment of depression following routine screening in patients with coronary heart disease or diabetes: a database cohort study. *Psychol Med* 2013;**43**:529–37. https://doi.org/10.1017/S0033291712001481

347. McLintock K, Russell AM, Alderson SL, West R, House A, Westerman K, Foy R. The effects of financial incentives for case finding for depression in patients with diabetes and coronary heart disease: interrupted time series analysis. *BMJ Open* 2014;**4**:e005178. https://doi.org/10.1136/bmjopen-2014-005178

348. Thombs BD, Ziegelstein RC. Does depression screening improve depression outcomes in primary care? *BMJ* 2014;**348**:g1253. https://doi.org/10.1136/bmj.g1253

349. O'Reilly SC, Muir KR, Doherty M. Knee pain and disability in the Nottingham community: association with poor health status and psychological distress. *Br J Rheumatol* 1998;**37**:870–3. https://doi.org/10.1093/rheumatology/37.8.870

350. Jinks C, Jordan K, Croft P. Measuring the population impact of knee pain and disability with the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC). *Pain* 2002;**100**:55–64. https://doi.org/10.1016/S0304-3959(02)00239-7

351. Jordan K, Jinks C, Croft P. A prospective study of the consulting behaviour of older people with knee pain. *Br J Gen Pract* 2006;**56**:269–76.

352. Dunn KM, Croft PR. Repeat assessment improves the prediction of prognosis in patients with low back pain in primary care. *Pain* 2006;**126**:10–15. https://doi.org/10.1016/j.pain.2006.06.005

353. Mallen CD, Thomas E, Belcher J, Rathod T, Croft P, Peat G. Point-of-care prognosis for common musculoskeletal pain in older adults. *JAMA Intern Med* 2013;**173**:1119–25. https://doi.org/10.1001/jamainternmed.2013.962

354. Keele CTU Quality Assurance Group. *SOP 17: Archiving and Destruction*. 2016. URL: www.keele.ac.uk/researchsupport/researchtoolkit/sops/healthandsocialcaresops/sop17/ (accessed 11 May 2017).

355. Jinks C, Ong BN, O'Neill TJ. The Keele community knee pain forum: action research to engage with stakeholders about the prevention of knee pain and disability. *BMC Musculoskelet Disord* 2009;**10**:85. https://doi.org/10.1186/1471-2474-10-85

356. Morden A, Jinks C, Ong BN. Lay models of self-management: how do people manage knee osteoarthritis in context? *Chronic Illn* 2011;**7**:185–200. https://doi.org/10.1177/1742395310391491

357. Brazier J, Roberts J, Tsuchiya A, Busschbach J. A comparison of the EQ-5D and SF-6D across seven patient groups. *Health Econ* 2004;**13**:873–84. https://doi.org/10.1002/hec.866

358. Al-Janabi H, Peters TJ, Brazier J, Bryan S, Flynn TN, Clemens S, *et al.* An investigation of the construct validity of the ICECAP-A capability measure. *Qual Life Res* 2013;**22**:1831–40. https://doi.org/10.1007/s11136-012-0293-5

359. Drummond MF, Sculpher MJ, Torrance GW, O'Brien B, Stoddart DL. *Methods for the Economic Evaluation of Health Care Programmes*. Oxford: Oxford University Press; 2005.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

360. van Hout BA, Al MJ, Gordon GS, Rutten FF. Costs, effects and C/E-ratios alongside a clinical trial. *Health Econ* 1994;**3**:309–19. https://doi.org/10.1002/hec.4730030505

361. Carpenter JR, Goldstein H, Kenward MG. REALCOM-IMPUTE software for multilevel multiple imputation with mixed response types. *J Stat Software* 2011;**45**.

Appendix 1 Patient and public involvement in the programme

Patient and public involvement in the Research Institute for Primary Care and Health Sciences, Keele University, prior to development of the National Institute for Health Research osteoarthritis programme

The research institute, IPCHS, at Keele involved lay people systematically in development, design, oversight and review of individual musculoskeletal research projects from the late 1990s.[120] In 2006, a RUG was formally established with the aim to embed PPI across the whole of its musculoskeletal programme and to encourage engagement in setting priorities for the programme. Twelve participants, who had a range of musculoskeletal conditions, including OA, were recruited from those who had been involved in earlier projects. A senior researcher from the centre led the RUG with administrative support and funding was provided. Activities of the RUG included formal discussion of new ideas and proposals with researchers, active representation in all centre projects and steering groups, presentations to funders, and organising a national conference for RUGs.

This was the broad background in place at the time that the idea for this programme grant was first mooted.

Patient and public involvement in development of the National Institute for Health Research osteoarthritis programme

The RUG members were involved in three important and specific ways, which directly influenced the development of the NIHR OA programme.

First, two members of RUG were playing an active role in the steering group of the institute's Knee Pain Prevention Project (KNEPP). This was a study, funded by the NIHR National Co-ordinating Centre for Research Capacity Development through its Public Health Initiative (principal investigator: Jinks), designed to investigate priorities for preventing disability in knee OA.[355] This study had set up quantitative and qualitative studies to carry out these investigations, together with a Community Knee Pain Forum,[22] which included patients and representatives from community and patient groups, voluntary agencies and health and social care practitioners. The two RUG members on the steering group were involved in discussion of the implications of the study findings. The evidence from the study was that self-management had substantial potential. The views of the RUG representatives were clear: people with knee OA understand the importance of self-management approaches (physical activity, weight control, self-medication with analgesia, searching out local resources), but they needed help to take this on, including sources of information about resources and advice about the details of an exercise programme. This input from the two RUG members had a major impact on shaping the theme of the NIHR programme, coming as it did at the same time as the idea of 'supported self-management' was emerging from research in other long-term conditions.[19] One of the two RUG members became a co-applicant on the funding proposal.

Second, and in parallel with KNEPP, one of the centre's researchers had been working closely with five RUG members on the development of a guidebook about OA for primary care patients, which had the aims of representing patient and health-care professional views and input into such information and of providing a guide and directory of local and national resources. Both of these aims contributed to the evolving idea of primary care providing support for self-management of OA. The guidebook differed from conventional patient education materials in that it contained lay as well as biomedical evidence-based knowledge. The RUG members reviewed a summary of qualitative research of people's experience of living with OA and combined this with their own experiences of living with OA as the basis for proposing

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

information for inclusion in the guidebook. They then reviewed a draft version. The guidebook development was written up as a research paper[142] by the lead researcher and the lead RUG member of the group and the guidebook itself was used as a central component of the MOSAICS intervention.

Third, RUG members had been actively involved in a successful NIHR Research for Patient Benefit award to IPCHS (principal investigator: Ong) to understand people's ideas about self-management of OA. This study had stemmed from the KNEPP study and in particular from the Community Knee Pain Forum discussions.[355] It was a qualitative study that engaged OA patients through in-depth interviews and contributed important information about the views and perceptions of OA patients in shaping and developing the OA programme. The study has been published.[22,176,356]

Patient and public involvement in Institute of Primary Care and Health Sciences during the programme

By 2009, the year in which the programme grant was awarded, the centre's portfolio of musculoskeletal research had grown substantially and in that year the RUG was expanded to 42 members. A dedicated PPI co-ordinator was appointed to replace the senior researcher who had been doing this in a part-time capacity. The new co-ordinator was a patient with chronic pain. Training and support was provided for RUG members, such as 'Contributing Assertively in Meetings', and RUG members undertook training of researchers in how to effectively involve lay people in their research. The formal activities of the RUG, all of direct relevance to the NIHR OA programme, now included prioritising research questions, devising effective and ethical recruitment strategies, interpreting findings and disseminating results.

By 2012, as the OA programme was entering its last phase, a PPI support worker, also a patient with chronic musculoskeletal problems, was appointed to work with the RUG co-ordinator. The RUG continued to be chaired by a senior academic from IPCHS during the period of the programme. This was perceived to be important from the perspective of the RUG members because it symbolised the extent of IPCHS commitment to PPI. From the institute's perspective, PPI was highlighted at strategic level and represented in senior management. This academic leadership ensured that all researchers understood that PPI was not optional, but an integral part of the way the programme operated. New researchers appointed to the programme learned about PPI as part of their induction. All research teams in the programme met with the PPI co-ordinator to plan and match RUG members with the requisite skills to the funded work.

The NIHR OA programme was taking place during this period of expansion of the PPI membership and infrastructure in IPCS, and the programme was one of the major areas of research in IPCHS that drew on and incorporated the RUG members and their expertise. An example of RUG involvement in individual projects in the programme is provided in *Patient and public involvement in workstream 2 (the MOSAICS studies)*.

Patient and public involvement in workstream 2 (the MOSAICS studies)

Two RUG members sat on the overall MOSAICS steering committee.

RUG collaboration on MOSAICS projects included participation in the Delphi consensus study on the content of the ideal OA consultation in primary care, engagement in the planning of the baseline population survey, membership of the group developing and writing the precise wording for quality indicators of OA care, analysis of the audio tapes of nurse-led OA clinics as part of one of the MOSAICS qualitative studies, and direct involvement in framing the content and style of training for GPs and health-care professionals for patients with OA.

The RUG involvement continues into the phase of disseminating the results of the programme to all people who have participated in the research projects. RUG members are helping to translate the results into

accessible language, including specific communications for control group participants who may feel they have missed out on opportunities for treatment.

The potential complexity of organising and supporting RUG contribution and involvement in the suite of studies in the MOSAICS workstream was managed by:

1. a glossary of terms used in the study literature
2. support for RUG members at meetings from the co-ordinator and support worker
3. lay summaries circulated in advance of meetings
4. continuing feedback to RUG members involved throughout the study
5. training.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Appendix 2 MOSAICS supporting documents

MOSAICS eligibility criteria

General practices and health-care professionals

- Member of the Central England PCRN or a Keele Research Network Practice.
- At least two GPs willing to undertake the study as per protocol (i.e. act as a control or intervention practice).
- Willing, and able, to allow one (or for preference two, to allow for cross-cover) of their practice nurses to be trained to deliver the nurse OA clinics.
- Able to physically accommodate the nurse-led clinics in the practice.
- Uses the EMIS computerised consultation system.
- Nurses and GPs consent to interview and follow-up by the MOSAICS studies team.
- GPs willing to be trained to carry out the GP OA consultations.
- Nurses willing to be trained to carry out the nurse OA clinics.
- Nurses consent to being observed and audio-recorded in their clinics.

Patients

Inclusion criteria

- Males and females.
- ≥ 45 years.
- Registered with a MOSAICS studies practice.
- For the individual-level analysis: pain in at least one joint site (hand, hip, knee or foot), consenting to further contact from the study team and medical record review (consent sought as part of the patient population survey).

Exclusion criteria

- Excluded via GP screen of practice list.
- Unable to give fully informed consent (e.g. learning difficulties or dementia).
- Resident in a care or nursing home.
- History of serious disease (e.g. malignancy, terminal illness).
- Unable to consult in the general practice surgery.
- Flagged as excluded from research in that practice.

MOSAICS model consultation task list

The GP model OA consultation: specific tasks promoted in the GP training workshops ordered by key model OA consultation tasks

Giving the diagnosis	
1.1	The GP elicits the patient's ideas or worries or concerns about what they think is the matter with them, or the cause of their problem
1.2	The GP tells the patient the problem is due to OA, the word OA needs to be used

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Explaining the diagnosis	
2.1	The GP elicits what the patient knows or understands about OA, the word OA needs to be used
2.2	The GP tells the patient that OA does not always/inevitably get worse, the word OA does NOT need to be used
2.3	The GP tells the patient that OA is treatable: that there are things which can be done to help, the word OA does NOT need to be used
Addressing expectations	
3.1	The GP elicits the specific expectation(s) the patient has of the GP about the problem
3.2	The GP responds to the patient's specific expectations (as noted at 3.1)
Providing analgesia	
4.1	The GP elicits what the patient has tried or is trying for the problem
4.2	The GP advises about, or prescribes for, pain relief
Promoting self-management	
5.1	The GP elicits what the patient has tried or is trying for the problem, other than for the pain
5.2	The GP tells the patient that exercise(s) or physical activity is beneficial for patients with OA or for the patient's problem
5.3	The GP tells the patient that losing weight, or not being overweight, is beneficial for patients with OA or for the patient's problem
Promoting self-management support	
6.1	The GP offers, or gives, the patient general written information on OA
6.2	The GP offers, or gives, the patient an appointment with a practice nurse to help with OA

Parts of this text have been reproduced with permission from Porcheret M, Main C, Croft P, Dziedzic K. Keele University. 2017.

Items on the MOSAICS e-template

Parts of this section have been reproduced from Edwards *et al*.[132] © The Author 2014. Published by Oxford University Press on behalf of the British Society for Rheumatology. This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com.

Quality indicator item	Possible response options	Quality indicator achieved if recorded as[a]
Pain assessment	None	None OR mild OR moderate OR severe
	Mild	
	Moderate	
	Severe	
Functional limitation assessment	None	None OR mild OR moderate OR severe
	Mild	
	Moderate	
	Severe	
Topical NSAID use	Tried full dose	Tried full dose OR offered full dose OR patient declined full dose OR not appropriate
	Offered full dose	
	Patient declined full dose	
	Not appropriate	
	Unknown	

Quality indicator item	Possible response options	Quality indicator achieved if recorded as[a]
Paracetamol use	Tried full dose	Tried full dose OR offered full dose OR patient declined full dose OR not appropriate
	Offered full dose	
	Patient declined full dose	
	Not appropriate	
	Unknown	
OA information given	Verbal and written[b]	Verbal and written OR verbal only OR not appropriate
	Verbal only	
	Not appropriate	
	Not this time	
Weight loss advice[c]	Verbal and written[b]	Verbal and written OR verbal only OR not appropriate
	Verbal only	
	Not appropriate	
	Not this time	
Exercise advice	Verbal and written[b]	Verbal and written OR verbal only OR not necessary or Not appropriate
	Verbal only	
	Not necessary	
	Not appropriate	
	Not this time	
Consideration of physiotherapy referral	Offered	Offered OR not necessary OR not appropriate
	Not necessary	
	Not appropriate	
	Not this time	

a Indicator not achieved if another of the possible response options used or no entry recorded.
b Response option used for assessment of effect of core NICE recommendations.
c In those with a recorded BMI of ≥ 25 kg/m^2 in the previous 3 years.

MOSAICS health economics methods

Parts of this section have been amended and reproduced from Oppong *et al.*[172] © The Author(s) 2018. Published by Oxford University Press on behalf of the British Society for Rheumatology. This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com.

Overview

The aim of the economic evaluation was to determine the cost-effectiveness of the MOAC in comparison with usual care in patients who consult with OA in primary care.

A cost–consequence analysis was initially conducted, describing all the important results relating to costs and consequences (EQ-5D, SF-6D, ICECAP-A). Subsequently, an incremental cost–utility analysis was then undertaken using patient responses to the EQ-5D-3L at baseline, 3, 6 and 12 months to estimate incremental costs, QALYs and net benefits. The base case was conducted from a NHS and PSS perspective.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Sensitivity analysis considered a wider societal perspective, for which private health-care costs and those associated with productivity loss were measured.

Resource use and cost data

Resource use data as a result of joint problems were collected primarily from the MOSAICS self-report postal questionnaires, which were completed by study participants at 6 and 12 months. NHS resource use data obtained included primary care contacts (e.g. GP and nurse visits), secondary care contacts (e.g. visits to the physiotherapist and occupational therapist) and medical investigations such as radiography and prescribed medication. Information on a participant's private expenditure included information on over-the-counter medication and visits to non-NHS health-care professionals. Information on employment status and time off work due to joint problems were also obtained from the participant questionnaire in order to estimate productivity loss over the 12-month period.

Resource use was valued by obtaining unit costs (2012/2013 prices) from standard sources such as PSSRU,[173] BNF[174] and NHS reference costs,[175] and applying them to resource use data. As a result of the lack of nationally representative unit cost estimates for private health care, this care was costed as the NHS equivalent. Costs associated with over-the-counter medications were obtained from patient self-report in the questionnaires or by allocating costs to the drugs mentioned in self-report. Unit costs of resource use items are presented in *Table 61*. Costs of the intervention were obtained in discussion with the study co-ordinators.

TABLE 61 Unit costs (£) of resource use items

Resource use	Unit cost (£)
Primary care visits[173]	
GP visits per 11.7 minutes	34
Nurse visits at practice per hour	44
Nurse visits at home per hour	60
Other healthcare professionals (attached to practice)	Participant specific
Secondary care visits[175]	
Orthopaedic surgeon	128
Orthopaedic surgeon (follow-up)	102
Physiotherapist	49
Physiotherapist (follow-up)	44
Occupational therapist	75
Occupational therapist (follow-up)	34
Podiatrist	74
Podiatrist (follow-up)	41
Other secondary care visits	Participant specific
Other resource use	
Private consultants	Costed to the NHS equivalent
Private other health-care professionals	Costed to the NHS equivalent
Hospital investigations/treatments	Participant specific
Prescribed drugs[174]	Participant specific
Over-the-counter drugs	Participant specific

Health outcome data

Information on health-related QoL was obtained from participant responses to the EQ-5D-3L questionnaire at baseline, 3, 6 and 12 months and the UK value set[253] was used to obtain EQ-5D index scores which were used to calculate QALYs over the 12-month period. Participants also completed the SF-12 questionnaire, which was used to generate SF-6D scores[357] and the ICECAP-A questionnaire[358] at baseline, 3, 6 and 12 months. The ICECAP-A is a measure of capability for adults, which aims to capture an individual's freedom to function in five key areas of their life: attachment, autonomy, enjoyment, stability and achievement.[358]

Data analysis

The data analysis was carried out on an intention-to-treat basis and focused on determining whether or not the MOAC is cost-effective when compared with usual care. Descriptive statistics were used to summarise the main health economic outcomes (EQ-5D-3L, SF-6D and ICECAP-A). Multiple imputation was used to account for missing EQ-5D, SF-6D, ICECAP-A scores at baseline 3, 6 and 12 months as well as missing costs at 6 and 12 months. Using the area under the curve approach, QALYs over 12 months were estimated for each study participant. Imbalances in baseline utility (EQ-5D scores) between the trial arms were controlled for using a regression approach.[257] Mean costs associated with each trial arm were estimated and, owing to the skewed nature of the costs, 95% CIs were calculated using non-parametric bootstrapping. Net monetary benefit, which is defined as the change in effectiveness (ΔE) multiplied by the cost-effectiveness threshold (λ) minus the change in cost (ΔC), was also estimated for each participant.[359] The threshold value (λ) used for the estimation of net benefits was £20,000 per QALY.

Multilevel modelling, an approach that has been recommended for the economic analysis of cluster trials, was used for the analysis of data. Dependent variables included net monetary benefits, costs, QALYs and cost of work absence while independent variables included sex and baseline EQ-5D. Model estimates of the difference in costs, QALYs and net monetary benefits were used to derive an incremental cost per QALY gained and an incremental net monetary benefit. Uncertainty was explored through the use of CEACs, which plot the probability that the intervention is cost-effective against willingness-to-pay threshold values.[360]

The human capital approach, which assumes that the value of lost work is equal to the amount of resources an individual would have been paid to do that work, was used to estimate productivity costs. The average wage for each respondent was identified using UK Standard Occupational Classification coding and annual earnings data for each job type and this was multiplied by self-reported days off work in order to estimate productivity costs.[258,259]

Sensitivity analysis focused on exploring uncertainties in the trial-based data by using QALYs generated from the SF-6D to obtain cost-effectiveness estimates and exploring broader costs through the inclusion of private health-care costs (e.g. over-the-counter medication costs and private health-care utilisation costs).

All analyses were carried out in Stata version 12, REALCOM-IMPUTE (Original version 2011; Centre for Multilevel Modelling, University of Bristol, Bristol, UK)[361] and Microsoft Excel® for Windows (Microsoft Office v12 2010; Microsoft Corporation, Redmond, WA, USA). Discounting was not required as the follow-up period was 12 months.

Results

A total of 525 participants were included in the study. Of these, 288 participants were randomised to the model OA consultation arm and 237 randomised to the usual-care arm. Follow-up rates at 6 and 12 months were 424 (81%) and 384 (73%), respectively. A total of 354 (67%) completed the postal questionnaire at both 6 and 12 months and a total of 305 (58%) participants provided complete EQ-5D data at all time points.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Resource use and costs

Resource use associated with the model OA consultation and usual-care interventions have been reported in *Table 62*. Most primary care resource use items were higher in the usual-care arm. The only exception was home visits by the GP, which were higher in the model OA consultation arm. With respect to secondary care, most resource use items were higher in the usual-care arm with the exception of visits to the orthopaedic surgeon which was significantly higher in the usual-care arm; all other secondary care items did not record any significant difference between trial arms. Approximately 59% of participants in the model OA consultation arm received a prescription compared with 65% in the usual-care arm, while approximately 49% of participants in the model OA consultation arm purchased over-the-counter medication for their joint problem compared with 46% in the usual-care arm.

Costs associated with resource use items are presented in *Table 63*. Both primary care and secondary care costs were higher in the usual-care arm and, overall, the model OA consultation arm was associated with a slightly lower cost (*Table 64*).

TABLE 62 Resource use mean (SD) over 12 months (complete cases)

Resource use category	Model OA consultation ($n = 199$)	Usual care ($n = 155$)	Difference (bootstrapped 95% CI)
Primary care visits[a]	1.52 (2.46)	1.99 (3.38)	–0.48 (–1.18 to 0.13)
GP at practice	1.32 (2.11)	1.59 (2.62)	–0.28 (–0.78 to 0.24)
GP at home	0.02 (0.12)	0.01 (0.08)	0.01 (–0.11 to 0.03)
Nurse at practice	0.19 (0.67)	0.39 (1.29)	–0.20 (–0.48 to –0.01)
Nurse at home	0	0.01 (0.08)	–0.01 (–0.03 to 0)
Other health-care professionals (attached to practice)[b]	0.21 (0.86)	0.32 (1.15)	–0.12 (–0.33 to 0.11)
Secondary care visits[c]	1.11 (2.65)	1.43 (2.91)	–0.32 (–0.96 to 0.27)
Orthopaedic surgeon	0.34 (0.89)	0.58 (1.37)	–0.24 (–0.52 to –0.003)
Podiatrist	0.13 (0.92)	0.12 (0.80)	0.003 (–0.17 to 0.17)
Physiotherapist	0.61 (2.01)	0.65 (1.93)	–0.04 (–0.47 to 0.36)
Occupational therapist	0.04 (0.21)	0.07 (0.58)	–0.04 (–0.16 to 0.04)
Other secondary care visits[b]	0.16 (0.91)	0.10 (0.51)	0.06 (–0.07 to 0.24)
Private consultants[d]	0.39 (1.66)	0.57 (3.07)	–0.18 (–0.79 to 0.29)
Private other health-care professionals[b]	0.13 (0.85)	0.04 (0.28)	0.09 (–0.02 to 0.23)
Hospital investigations/treatments[b,e]	82 (41.21%)	72 (46.45%)	10
Prescribed drugs[b,e]	117 (58.79%)	101 (65.16%)	16
Over-the-counter drugs[b,e]	98 (49.25%)	72 (46.45%)	26

a Includes contacts with GP and nurse at home and practice.
b Patient specific.
c Includes contacts with physiotherapists, occupational therapists, etc.
d Includes contacts with private physiotherapists, occupational therapists, etc.
e The proportion of patients who reported that they had undergone a hospital investigation for their OA during follow-up and the proportion who reported that they had been prescribed medication for their OA (presented separately for 'prescribed' and 'over the counter').

TABLE 63 Mean costs (SD) per patient over 12 months (complete cases) for each resource category in OA care

Resource use category	Model OA consultation ($n = 199$) (£)	UC ($n = 155$) (£)	Difference (bootstrapped 95% CI)
Primary care visits[a]	56.01 (83.53)	69.02 (103.31)	–13.01 (–35.24 to 5.28)
GP at practice	44.76 (71.80)	54.18 (89.01)	–9.42 (–29.03 to 7.41)
GP at home	0.81 (6.55)	0.35 (4.31)	0.46 (–0.71 to 1.55)
Nurse at practice	2.11 (7.46)	4.61 (15.07)	–2.49 (–5.50 to –0.03)
Nurse at home	0	0.15 (1.87)	–0.15 (–0.54 to 0)
Other primary care visits[b]	8.33 (24.20)	9.74 (29.72)	–1.41 (–7.37 to 3.97)
Secondary care visits[c]	60.68 (130.42)	76.48 (156.38)	–15.80 (–51.40 to 14.01)
Orthopaedic surgeon	27.09 (71.66)	44.31 (106.94)	–17.22 (–37.95 to 1.18)
Podiatrist	5.32 (35.65)	4.34 (26.23)	0.98 (–5.03 to 7.93)
Physiotherapist	21.55 (77.01)	21.74 (70.72)	–0.18 (–15.47 to 16.50)
Occupational therapist	2.06 (11.98)	2.24 (16.93)	–0.18 (–3.50 to 2.61)
Other secondary care visits[b]	4.67 (17.85)	3.85 (22.67)	0.81 (–4.45 to 4.65)
Hospital investigations/treatments[b]	109.71 (401.16)	92.36 (222.66)	17.35 (–42.40 to 83.75)
Prescribed drugs[b]	15.51 (20.34)	15.65 (21.47)	–0.14 (–4.58 to 3.86)
Trial intervention cost	11.47 (20.69)	0	11.47 (8.69 to 14.42)
Over-the-counter drugs[b]	27.14 (255.67)	27.93 (121.01)	–0.79 (–31.51 to 50.14)
Private health professionals[b]	21.62 (76.54)	29.53 (135.05)	–7.91 (–39.24 to 12.24)

a Includes contacts with GP and Nurse at home and practice.
b Patient-specific.
c Includes contacts with physiotherapists, occupational therapists etc.

TABLE 64 Mean total costs (SD) per patient over 12 months (imputed analysis)

Health-care setting	Model OA consultation ($n = 288$)	Usual care ($n = 237$)	Difference (bootstrapped 95% CI)
Total NHS costs[a]	227.17 (411.84)	236.11 (345.35)	–8.94 (–71.79 to 57.70)
Total Healthcare costs[a]	278.56 (535.43)	285.99 (400.43)	–7.43 (–76.41 to 76.26)

a Unadjusted costs.

Health outcomes

The EQ-5D and SF-6D scores increased at all time points over the 12-month period in both the model OA consultation and usual-care arms, indicating an improvement in health status over time. Although these scores were higher in the usual-care arm, the differences were not statistically significant. Compared with baseline, ICECAP-A scores in the model OA consultation arm were higher at 12 months, indicating an increase in capability over time, while the opposite was true for the usual-care arm for which ICECAP-A scores at 12 months were lower than the baseline value. When QALYs were estimated, the usual-care arm was associated with a marginally higher total QALYs (*Table 65*).

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

TABLE 65 Health outcomes, mean (SD) over 12 months (imputed analysis)

Outcome measure and time of measurement where relevant	Model OA consultation ($n = 288$)	Usual care ($n = 237$)	Difference (bootstrapped 95% CI)
EQ-5D scores			
Baseline	0.573 (0.298)	0.588 (0.272)	–0.015 (–0.062 to 0.039)
Month 3	0.615 (0.280)	0.631 (0.264)	–0.016 (–0.064 to 0.030)
Month 6	0.637 (0.264)	0.638 (0.259)	–0.001 (–0.044 to 0.044)
Month 12	0.651 (0.262)	0.674 (0.224)	–0.023 (–0.067 to 0.018)
QALYs	0.627 (0.244)	0.639 (0.224)	–0.012 (–0.054 to 0.026)
QALYs[a]	0.632	0.634	–0.002
QALYs[b]			–0.003 (–0.026 to 0.197)
SF-6D scores			
Baseline	0.678 (0.139)	0.690 (0.148)	–0.012 (–0.037 to 0.013)
Month 3	0.688 (0.141)	0.696 (0.141)	–0.008 (–0.033 to 0.017)
Month 6	0.687 (0.142)	0.707 (0.144)	–0.020 (–0.044 to 0.004)
Month 12	0.693 (0.139)	0.702 (0.138)	–0.009 (–0.032 to 0.015)
QALYs	0.688 (0.128)	0.701 (0.129)	–0.013 (–0.038 to 0.010)
QALYs[a]	0.692	0.696	–0.004
QALYs[b]			–0.012 (–0.03 to 0.01)
ICECAP-A			
Baseline	0.826 (0.166)	0.851 (0.155)	–0.025 (–0.053 to 0.003)
Month 3	0.828 (0.151)	0.853 (0.155)	–0.025 (–0.053 to 0.001)
Month 6	0.821 (0.160)	0.843 (0.158)	–0.022 (–0.049 to 0.005)
Month 12	0.837 (0.153)	0.846 (0.155)	–0.009 (–0.038 to 0.014)

a Adjusted for baseline utility.
b Adjusted for baseline utility and sex (regression model).

Cost-effectiveness

The results from the base-case cost-effectiveness analysis showed that the model OA consultation was associated with a smaller cost ($p = 0.705$) and is not as effective ($p = 0.786$) as usual care. However, the difference was not statistically significant (*Table 66*). Net monetary benefit at £20,000 per QALY was negative ($p = 0.887$) and the CEACs showed that, at a willingness-to-pay threshold of £20,000 per QALY, there is a 44% chance of the model OA consultation being cost-effective (*Figure 16*).

TABLE 66 Base-case cost–utility analysis (imputed analysis)

Cost or utility measure	Difference (intervention–control)	*p*-value	CI	Interpretation
NHS costs[a]	–13.11	0.705	–81.09 to 54.85	Intervention less costly and less effective. Positive ICER but negative net benefits
QALYs[a]	–0.003	0.786	–0.03 to 0.02	
Net monetary benefits[a]	–33.63	0.887	–497.56 to 430.30	

a Adjusted for baseline utility and sex (three-level model).

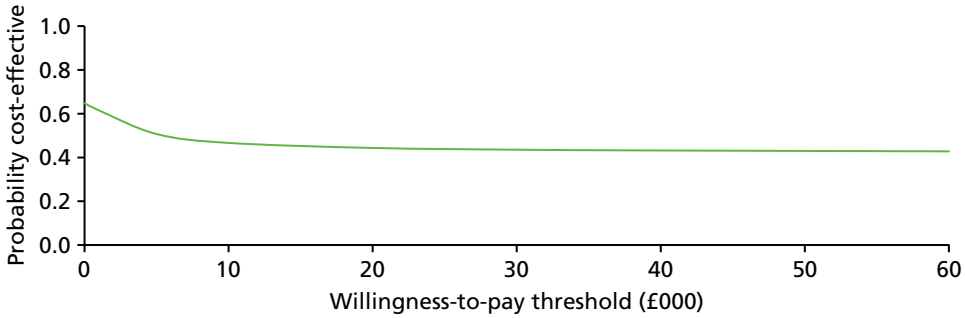

FIGURE 16 Base-case cost–utility analysis (imputed analysis): CEAC.

Sensitivity analysis with broader health-care costs indicated that the intervention was still less costly ($p = 0.768$) and less effective ($p = 0.786$) than the usual-care arm and when the SF-6D was used as an outcome measure, a similar result to the base-case analysis was achieved. Cost–utility analysis with QALYs generated from the SF-6D yielded similar results to the base-case analysis, that is, the intervention was less costly ($p = 0.705$) and less effective ($p = 0.187$) than usual care (*Table 67*).

Work-related outcomes

Participants in the intervention arm had fewer mean days off work than those in the usual-care arm ($p = 0.364$). The associated productivity related cost was lower in the intervention arm but the difference was not statistically significant (see *Table 67*).

Conclusion

In summary, implementing NICE OA guidelines[15] in primary care using the MOSAICS intervention does not lead to increased costs and appears to reduce demand for orthopaedic surgery and time lost from work. However, QoL was barely changed.

TABLE 67 Sensitivity analysis

Cost or utility measure	Difference (intervention–control)	*p*-value	CI	Interpretation
Cost-utility analysis with SF-6D				
NHS costs[a]	–13.11	0.705	–81.09 to 54.85	Intervention less costly and less effective. Positive ICER but negative net benefits
QALYs (SF-6D)[a]	–0.012	0.187	–0.03 to 0.01	
Net monetary benefits[a]	–178.39	0.362	–561.74 to 204.96	
Cost-utility analysis with health-care costs				
Health-care costs[a]	–14.14	0.768	–108.08 to 79.80	Intervention less costly and less effective; thus, the usual care dominates
QALYs[a]	–0.003	0.786	–0.03 to 0.02	
Net monetary benefits[a]	–34.95	0.883	–501.82 to 431.92	
Time off work and productivity costs				
Number of days off over 12 months[a]	–1.05	0.364	–3.35 to 1.23	
Mean cost (£) of work absence[a]	–23.25	0.845	–256.32 to 209.83	

a Adjusted for baseline utility, and sex (regression model).

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Appendix 3 BEEP trial supporting documents

Interview guides (post intervention)

The BEEP interview Study

Interview Guide 1 (Version 1, 12/01/10)

Probes/ prompts shown in italics

PRELIMINARY POINTS

- **Check participant's personal details**
- **Explain purpose of study; researcher to introduce self; confirm confidentiality**
- **Introduction – interested in participant experience of physiotherapy treatment received, particularly the exercise programme**
- **No right or wrong answers**

COMPLETE CONSENT FORM A

Introduction:

1. Can you tell me a little bit about your knee problem?

2. Have you had any previous treatments for your knee problem?

   E.*g. from GP/ previous physiotherapy/ previous exercise programme*

3. What do you think the future holds for your knee problem?

4. What did you expect from your physiotherapy treatment?

   Treatment expectation/ outcome expectation

Main interview:

5. Can you tell me what happened during your physiotherapy sessions?

   Exercise/ other treatment/ advice/ number of sessions

6. Can you tell me about your treatment?

   The exercise programme: type of exercises/ setting of exercise

7. How was *that* exercise programme decided upon (i.e. that type of exercise, that setting of exercise)?

   Patient input/ physiotherapist decision/ acceptable

8. Overall, how do you feel you got on with the exercises that you were asked to do?

   Did you do the exercises as often as you were asked/ Are you still doing the exercises now/ Do you think you will carry them on in the future? Why?

9. Have you got any plans for the future with regards to treatment?

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Possibility of return to physiotherapy/ onward referral/ how was this decision made/ acceptable?

10. Do you think that your treatment helped?

How/ why/ why not?

11. Was anything particularly good (and bad) about your experience?

12. Is there anything that you would have liked to have happened differently?

Conclusion:

13. Is there anything else that you would like to tell me about your experience of physiotherapy and exercise?

COMPLETE CONSENT FORM B

End of interview

Interview guide (18 months follow-up)

The BEEP interview Study

Interview Guide 2 (Version 1, 12/01/10)

Probes/ prompts shown in italics

PRELIMINARY POINTS

- **Check participant's personal details**
- **Explain purpose of study; researcher to introduce self; confirm confidentiality**
- **Introduction – interested in participant experience of physiotherapy treatment received, particularly the exercise programme**
- **No right or wrong answers**

COMPLETE CONSENT FORM A

Introduction:

1. Can you tell me a little bit about your knee problem?

2. Has your knee problem changed since your last interview?

2. Have you had any more treatment for your knee pain, since your physiotherapy sessions?

E.*g. from GP/ more physiotherapy/tried things yourself at home/ why?*

3. What do you think the future holds for your knee problem?

Main interview:

4. Can you tell me a little bit about the physiotherapy you received?

Exercise/ other treatment/ advice

5. Did you feel that your treatment has helped?

How /why/ why not?

6. Overall, how do you feel you got on with the exercises that you were asked to do?

Did you do the exercises as often as you were asked/ Are you still doing the exercises now/ Do you think you will carry them on in the future? Why?

7. Was anything particularly good (and bad) about your experience?

8. Is there anything that you would have liked to have happened differently?

Conclusion:

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

9. Is there anything else that you would like to tell me about your experience of physiotherapy and being given an exercise programme?

COMPLETE CONSENT FORM B

End of interview

Appendix 4 Related publications

Commentaries, editorials or reviews by programme team members linked to work of the programme

Blagojevic M, Jinks C, Jeffery A, Jordan KP. Risk factors for onset of osteoarthritis of the knee in older adults: a systematic review and meta-analysis. *Osteoarthr Cartil* 2010;**18**:24–33.

Cottrell E, Roddy E, Foster NE. The attitudes, beliefs and behaviours of GPs regarding exercise for chronic knee pain: a systematic review. *BMC Fam Pract* 2010;**11**:4.

Croft P, Porcheret M, Peat G. Managing osteoarthritis in primary care: the GP as public health physician and surgical gatekeeper. *Br J Gen Pract* 2011;**61**:485–6.

Dziedzic KS, Hill JC, Porcheret M, Croft PR. New models for primary care are needed for osteoarthritis. *Phys Ther* 2009;**89**:1371–8.

Dziedzic KS, Healey EL, Main CJ. Implementing the NICE osteoarthritis guidelines in primary care: a role for practice nurses. *Musculoskeletal Care* 2013;**11**:1–2.

Dziedzic KS. Osteoarthritis: best evidence for best therapies in hand osteoarthritis. *Nat Rev Rheumatol* 2011;**7**:258–60.

Fernandes L, Hagen KB, Bijlsma JW, Andreassen O, Christensen P, Conaghan PG, *et al.* European League Against Rheumatism (EULAR). EULAR recommendations for the non-pharmacological core management of hip and knee osteoarthritis. *Ann Rheum Dis* 2013;**72**:1125–35.

Foster NE, Dziedzic KS, van der Windt DA, Fritz JM, Hay EM. Research priorities for non-pharmacological therapies for common musculoskeletal problems: nationally and internationally agreed recommendations. *BMC Musculoskelet Disord* 2009;**10**:3.

Foster NE, Hartvigsen J, Croft PR. Taking responsibility for the early assessment and treatment of patients with musculoskeletal pain: a review and critical analysis. *Arthritis Res Ther* 2012;**14**:205.

Helliwell T, Mallen C, Peat G, Hay E. Research into practice: improving musculoskeletal care in general practice. *Br J Gen Pract* 2014;**64**:372–4.

Lau R, Stevenson F, Ong BN, Dziedzic K, Eldridge S, Everitt H, *et al.* Addressing the evidence to practice gap for complex interventions in primary care: a systematic review of reviews protocol. *BMJ Open* 2014;**4**:e005548.

Mallen CD. Osteoarthritis – the forgotten chronic disease. Time for a multimorbid approach? *Eur J Gen Pract* 2012;**18**:131–2.

Morden A, Jinks C, Ong BN. Understanding Help Seeking for Chronic Joint Pain: Implications for Providing Supported Self-Management. *Qual Health Res* 2014;**24**:957–68.

Ong BN, Rogers A, Kennedy A, Bower P, Sanders T, Morden A, *et al.* Behaviour change and social blinkers? The role of sociology in trials of self-management behaviour in chronic conditions. *Sociol Health Illn* 2014;**36**:226–38.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Peat G, Birrell F, Cumming J, Doherty M, Simpson H, Conaghan PG, Arthritis Research UK Clinical Studies Group for Osteoarthritis and Crystal Diseases. Under-representation of the elderly in osteoarthritis clinical trials. *Rheumatology* 2011;**50**:1184–6.

Porcheret M, Healey E, Dziedzic KS. Uptake of best arthritis practice in primary care – no quick fixes. *J Rheumatol* 2011;**38**:791–3.

Sanders T, Ong BN, Sowden G, Foster N. Implementing change in physiotherapy: professions, contexts and interventions. *J Health Organ Manag* 2014;**28**:96–114.

Silverwood V, Blagojevic-Bucknall M, Jinks C, Jordan JL, Protheroe J, Jordan KP. Current evidence on risk factors for knee osteoarthritis in older adults: a systematic review and meta-analysis. *Osteoarthr Cartil* 2015;**23**:507–15.

Strauss VY, Carter P, Ong BN, Bedson J, Jordan KP, Jinks C, Arthritis Research UK Research Users' Group. Public priorities for joint pain research: results from a general population survey. *Rheumatology* 2012;**51**:2075–82. https://doi.org/10.1093/rheumatology/kes179

Thomas E, Peat G, Croft P. Defining and mapping the person with osteoarthritis for population studies and public health. *Rheumatology* 2014;**53**:338–45.

Zhang W, Nuki G, Moskowitz RW, Abramson S, Altman RD, Arden NK, *et al.* OARSI recommendations for the management of hip and knee osteoarthritis: part III: Changes in evidence following systematic cumulative update of research published through January 2009. *Osteoarthr Cartil* 2010;**18**:476–99.

Zhang W, Doherty M, Peat G, Bierma-Zeinstra MA, Arden NK, Bresnihan B, *et al.* EULAR evidence-based recommendations for the diagnosis of knee osteoarthritis. *Ann Rheum Dis* 2010;**69**:483–9.

Linked studies relevant to the programme published by programme team members

Bedson J, Kadam U, Muller S, Peat G. Does comorbid disease influence consultation for knee problems in primary care? *Prim Health Care Res Dev* 2011;**12**:322–8.

Bedson J, Belcher J, Martino OI, Ndlovu M, Rathod T, Walters K, *et al.* The effectiveness of national guidance in changing analgesic prescribing in primary care from 2002 to 2009: an observational database study. *Eur J Pain* 2013;**17**:434–43.

Cheraghi-Sohi S, Bower P, Kennedy A, Morden A, Rogers A, Richardson J, *et al.* Patient priorities in osteoarthritis and comorbid conditions: a secondary analysis of qualitative data. *Arthritis Care Res* 2013;**65**:920–7.

Clarson LE, Nicholl BI, Bishop A, Edwards JJ, Daniel R, Mallen CD. Monitoring Osteoarthritis: A Cross-sectional Survey in General Practice. *Clin Med Insights Arthritis Musculoskelet Disord* 2013;**6**:85–91.

Hay AD, Rortveit G, Purdy S, Adams J, Sanci LA, Schermer TR, *et al.* Primary care research – an international responsibility. *Fam Pract* 2012;**29**:499–500.

Henschke N, Ostelo RW, Terwee CB, van der Windt DA. Identifying generic predictors of outcome in patients presenting to primary care with nonspinal musculoskeletal pain. *Arthritis Care Res* 2012;**64**:1217–24.

Holden MA, Nicholls EE, Young J, Hay EM, Foster NE. Exercise and physical activity in older adults with knee pain: a mixed methods study. *Rheumatology* 2015;**54**:413–23.

Holden MA, Nicholls EE, Young J, Hay EM, Foster NE. Role of exercise for knee pain: what do older adults in the community think? *Arthritis Care Res* 2012;**64**:1554–64.

Jinks C, Vohora K, Young J, Handy J, Porcheret M, Jordan KP. Inequalities in primary care management of knee pain and disability in older adults: an observational cohort study. *Rheumatology* 2011;**50**:1869–78.

Jordan KP, Jöud A, Bergknut C, Croft P, Edwards JJ, Peat G, *et al.* International comparisons of the consultation prevalence of musculoskeletal conditions using population-based healthcare data from England and Sweden. *Ann Rheum Dis* 2014;**73**:212–18.

Morden A, Jinks C, Ong BN. Rethinking 'risk' and self-management for chronic illness. *Soc Theory Health* 2012;**10**:78–99.

Morden A, Jinks C, Ong BN. '..I've found once the weight had gone off, i've had a few twinges, but nothing like before'. Exploring weight and self-management of knee pain. *Musculoskeletal Care* 2014;**12**:63–73.

Ong BN, Jinks C, Morden A. The hard work of self-management: Living with chronic knee pain. *Int J Qual Stud Health Well-being* 2011;**6**:7035.

Peat G, Thomas E, Duncan R, Wood L. Is a "false-positive" clinical diagnosis of knee osteoarthritis just the early diagnosis of pre-radiographic disease? *Arthritis Care Res* 2010;**62**:1502–6.

Plüddemann A, Wallace E, Bankhead C, Keogh C, Van der Windt D, Lasserson D, *et al.* Clinical prediction rules in practice: review of clinical guidelines and survey of GPs. *Br J Gen Pract* 2014;**64**:e233–42.

Sheikh L, Nicholl BI, Green DJ, Bedson J, Peat G. Osteoarthritis and the rule of halves. *Osteoarthr Cartil* 2014;**22**:535–9.

Thomas MJ, Moore A, Roddy E, Peat G. "Somebody to say 'come on we can sort this'": a qualitative study of primary care consultation among older adults with symptomatic foot osteoarthritis. *Arthritis Care Res* 2013;**65**:2051–5.

Background publications informing the programme

Grime J, Richardson JC, Ong BN. Perceptions of joint pain and feeling well in older people who reported being healthy: a qualitative study. *Br J Gen Pract* 2010;**60**:597–603.

Jinks C, Ong BN, O'Neill T. "Well, it's nobody's responsibility but my own." A qualitative study to explore views about the determinants of health and prevention of knee pain in older adults. *BMC Public Health* 2010;**10**:148.

Jinks C, Ong BN, O'Neill TJ. The Keele community knee pain forum: action research to engage with stakeholders about the prevention of knee pain and disability. *BMC Musculoskelet Disord* 2009;**10**:85.

Jordan KP, Wilkie R, Muller S, Myers H, Nicholls E, Arthritis Research Campaign National Primary Care Centre. Measurement of change in function and disability in osteoarthritis: current approaches and future challenges. *Curr Opin Rheumatol* 2009;**21**:525–30.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

May CR, Mair F, Finch T, MacFarlane A, Dowrick C, Treweek S, *et al.* Development of a theory of implementation and integration: Normalization Process Theory. *Implement Sci* 2009;**4**:29.

Murray E, Treweek S, Pope C, MacFarlane A, Ballini L, Dowrick C, *et al.* Normalisation process theory: a framework for developing, evaluating and implementing complex interventions. *BMC Med* 2010;**8**:63.

Appendix 5 Other grants linked to this programme

Arthritis Research Campaign Primary Care Musculoskeletal Research Centre of Excellence. Funder reference: 18139.

Professor Pauline Ong: NIHR Research for Patient Benefit. Grant reference: PB- PG-0107-11221.

Dr Mel Holden: Arthritis Research Campaign Allied Health Professionals Training Fellowship. Funder reference: 18004.

Dr John Edwards NIHR In-practice Fellowship. Grant reference: IAT/I-PF/010/009.

Dr Mark Porcheret NIHR In-practice Fellowship. Grant reference: 2008/IPF/PORCHERET.

Dr Andrew Finney NIHR Clinical Doctoral Research Fellowship. Grant reference: CAT CDRF 10-018.

Robert Smith, Keele University. ACORN* PhD studentship.

Dr Jerome Wulff, Keele University. ACORN* PhD studentship.

*ACORN is the name of the PhD Fellowship scheme funded by Keele University.

© Queen's Printer and Controller of HMSO 2018. This work was produced by Hay *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

EME
HS&DR
HTA
PGfAR
PHR

Part of the NIHR Journals Library
www.journalslibrary.nihr.ac.uk

This report presents independent research funded by the National Institute for Health Research (NIHR). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health

Published by the NIHR Journals Library